Sexual network structure and the spread of HIV in Africa: evidence from Likoma Island, Malawi

Stéphane Helleringer and Hans-Peter Kohler

Background: Whereas sexual relationships among low-risk individuals account for the majority of HIV infections in sub-Saharan Africa, limited knowledge exists about the structure and characteristics of sexual networks among the general population in sub-Saharan Africa.

Objectives: To investigate the population-level structure of sexual networks connecting the young adult population of several villages on Likoma Island (Malawi), and analyse the structural position of HIV-positive individuals within the sexual network.

Design and methods: A cross-sectional sociocentric survey of sexual partnerships and biomarkers of prevalent HIV infections.

Results: The study documents the existence of a large and robust sexual network linking a substantial fraction of the island's young adult population: half of all sexually active respondents were connected in a giant network component, and more than a quarter were linked through multiple independent chains of sexual relationships. This high network connectivity emerges within short time frames. The prevalence of HIV also varied significantly across the network, with sparser regions having a higher HIV prevalence than densely connected components. Several risk factors related to sexual mixing patterns help explain differentials in HIV prevalence across network locations.

Conclusion: Contrary to claims that sexual networks in rural sub-Saharan Africa are too sparse to sustain generalized HIV epidemics, the structure of the networks observed in Likoma appears compatible with a broad diffusion of HIV among lower-risk groups. The non-homogeneous distribution of HIV infection within the network suggests that network characteristics are an important determinant of the dynamics of HIV spread within a population.
© 2007 Wolters Kluwer Health | Lippincott Williams & Wilkins

AIDS 2007, **21**:2323–2332

Keywords: HIV risk factors, HIV transmission, Malawi, sexual behavior, sexual networks

Introduction

Studies of sexual networks in sub-Saharan Africa have focused primarily on spread networks that connect core group members (e.g. sex workers) with members of bridge populations (e.g. clients of sex workers) or the general population [1–5]. In generalized HIV epidemics such as those of southern and eastern Africa, however, infections within maintenance networks, that is networks of sexual relationships among low-risk individuals, account for the majority of new HIV cases [6]. Despite the importance of sexual networks among low-risk individuals [7–9], available evidence about their structure is limited. In particular, existing studies are exclusively based on a set of egocentric surveys that do not allow direct observation of the complete population-level structure of a sexual network in the general population [10]. Therefore, only a limited number of characteristics of sexual networks have been empirically investigated in sub-Saharan Africa, including for example age-mixing [11,12], spatial bridging [13–15], concurrency of sexual partnerships [16,17], and heterogeneity in rates of partner change [18–21].

To overcome this limitation and investigate, for the first time in sub-Saharan Africa, the structure of sexual

From the University of Pennsylvania, Population Studies Center, Philadelphia, Pennsylvania, USA.

Correspondence to Hans-Peter Kohler, University of Pennsylvania, Population Studies Center, 3718 Locust Walk, Philadelphia, PA 19104-6299, USA.

E-mail: hpkohler@pop.upenn.edu
Received: 10 October 2006; revised: 4 April 2007; accepted: 11 April 2007.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

networks in the general population and the position of HIV-positive individuals within these networks, we conducted a cross-sectional sociocentric survey of sexual partnerships among young adults in several villages of Likoma, an island on Lake Malawi. Similar surveys have been conducted elsewhere [22–25], but except for Bearman *et al.* [25], they have mostly been conducted among high-risk groups or in the context of concentrated epidemics.

Methods

The data collection for our sociocentric study of sexual partnerships occurred in two stages. First, we conducted a census of every individual on Likoma Island to obtain a roster of potential partners. Second, we conducted a sexual network survey with all inhabitants aged 18–35 years of seven geographically contiguous villages (Fig. 1), asking respondents for information about their romantic and sexual partners. The saturated sampling frame used in this study then allowed us to construct the population-level sexual network by matching the reported sexual partners with the census roster, and then linking the data of all young adults residing in the sample villages. The context and methodology of this survey are summarized below, and additional details are provided in Helleringer *et al.* [26]. The 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 before interviews and HIV testing.

Incomplete network bias affecting sociocentric network designs implies that the obtained networks are merely 'quasi-complete' [10,27]. This bias usually occurs because: (i) some network members reside outside of the research area and cannot be enrolled; (ii) adequate information is not always available to link all reported network partners; and (iii) respondents often do not report all of their sexual partnerships [10]. We used audio computer-assisted self-interview (ACASI) to minimize the misreporting of sexual behaviors [28,29]. The island setting of Likoma was chosen to reduce the potential biases from the other two sources. Likoma extends over only 18 square kilometers, has limited transportation to the mainland, and its population is small, with just over 7000 individuals living in a dozen villages (Fig. 1). As a result, and essential for the data quality of this project, a limited set of identifying information allowed the matching of nominated partners with the census rosters.

Data collection

Roster of potential sexual partners

All households in Likoma and Chizumulu (a neighboring island 5 km from Likoma) were enumerated to establish a

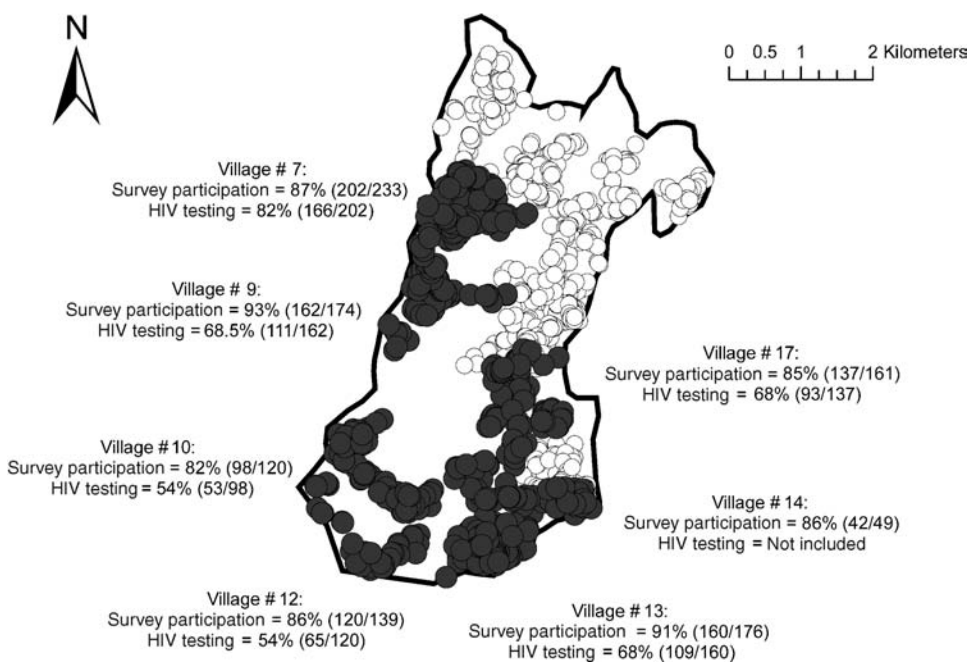

Fig. 1. Geographical location of the sampled villages and village-specific participation rates. Each circle represents a dwelling unit. Dark circles represent dwelling units in the villages that were included in the sexual network survey. Empty circles represent dwelling units in the villages that were not included in this sampling frame. Denominators of the survey participation rates are the total number of eligible respondents (aged 18–35 years 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 the location of dwelling units are approximate.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

detailed roster of potential sexual partners. Household informants were asked to provide names, maiden names (for married women), nicknames, and sociodemographic characteristics of all household residents. More than 1300 households were listed in Likoma, and approximately 500 in Chizumulu. The enumeration also included: (i) temporary migrants (i.e. household residents who were temporarily absent); (ii) family members who had moved permanently during the past 5 years; and (iii) family members who had died during the past 5 years.

Sexual network survey

We subsequently conducted an ACASI sexual network survey with all inhabitants aged 18–35 years (and their spouses) in the sampled villages. A total of 923 participants (501 women and 422 men) were interviewed. The age range of 18–35 years was selected because most HIV infections occur during that span [30]. The location of sampled villages, as well as participation rates, are displayed in Fig. 1. Initially, the two villages with the highest proportion of extramarital births were selected as seed villages (villages 13 and 9), and the remaining five villages were selected based on contiguity to these seeds. The names of up to five partners with whom a respondent had been in a romantic relationship within 3 years before the survey were recorded using headsets. The recall period of 3 years was chosen because of the long duration of HIV infectivity. A total of 2040 reports of partnerships were collected during this network survey, among which 1858 (91.0%) involved sexual intercourse. If a nominated partner was living in Likoma (currently or at the time of the relationship), full names were collected (see Helleringer *et al.* [26] for a further discussion of this name generator), and respondents were asked to specify his/her residence within the island (village, proximity to a landmark). Using this information, we attempted to identify each nominated partner within the roster of potential partners. This partner tracing was initially conducted using phonetic name-matching algorithms in STATA [31], and the links were then verified or completed by manual inspection. More than 80% of sexual relationships with a partner residing in Likoma were successfully linked to a record in the roster obtained from the island census [26]. A sexual relationship was assumed to exist if it was reported by at least one partner.

Partnership and health measures

The sexual network survey instrument included, for each reported relationship, the starting and end dates, the location and occasion of original encounter, the ways in which partners knew each other before the start of the relationship, whether the relationship involved sexual intercourse, the frequency of sexual intercourse, condom use during the relationship, and other measures of relationship quality. All answers to these questions were categorical. In particular, dates of relationships were recorded using the following categories: more than a year ago/within past year/within past month/ongoing. Non-marital partnerships were classified as either stable relationships, infrequent partnerships or one-off casual encounters. Condom use was defined as ‘consistent’ when the respondent reported always using condoms with all partners. Respondents were also asked about their own health and healthcare utilization, including the presence of sexually transmitted infection (STI) symptoms (indicating the presence of either ulceration of the genital area, discharge or inflamation/irritation during urination), having ever been tested for HIV, and having received an injection during the year before the survey. Finally, respondents in six villages were tested for HIV ($N = 597$; the seventh village, village 14 in Fig. 1, could not be included because of funding and timing constraints). HIV serostatus was determined using two rapid test assays (Determine HIV/1-2 and UniGold HIV), with participation rates varying across villages from 54 to 82% ($P < 0.01$; Fig. 1).

Graph-theoretical measures are used to describe network properties [32]. As a person A may transmit infection to B if there is a path of sexual relations (indirectly) connecting them, a set of individuals in which each pair is connected by at least one path (a component), is an important epidemiological unit. The distribution of component sizes limits the extent to which a disease can spread within a population. Epidemic transmission of a pathogen is also associated with cycles, that is, paths that begin and end at the same person. For example, consider A, B, C and D are connected linearly as A–B–C–D. If A fails to infect B, then A has no chance of infecting C. If they are connected cyclically (e.g. A–B–C–D–A), then if A fails to infect B, A can still reach C through D. Cyclical networks are robust because they remain connected even if some individuals or relationships are removed from the network. We measure network robustness by using bicomponents, that is subsets of individuals that are connected by at least two independent paths within the sexual network and are thus potentially at an increased risk of HIV infection. Network graphs were produced using PAJEK [33].

The duration, timing and overlaps of partnerships are additional important determinants of epidemic dynamics [16]. For example, if A and B are in a partnership at time 1, and A and C are in a partnership at time 2, then B can infect C but C cannot infect B. If both partnerships are ongoing at the same time, on the other hand, infection can reach everyone in this network [16,34]. Because our measures of partnership timing and duration are relatively imprecise, we cannot reliably assess whether partnerships occur concurrently, or follow each other serially. We therefore proxy the evolution of the network over time by comparing ‘snapshots’ of the network of active relationships during various time periods (for a similar approach using prospective data, see Rothenberg *et al.* [23]).

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Results

Respondents characteristics

Table 1 describes the characteristics of the 923 survey respondents. The mean age for women in the sample was 24.6 years [interquartile range (IQR) 20–28] and 26 years (IQR 21–30) for men. Approximately two thirds of the sample were born in Likoma. A total of 597 participants were tested for HIV infection and 48 cases were identified. The prevalence of HIV infection in the sample was higher among women [10.6%, 95% confidence interval (CI) 7.2–13.9] than among men (4.7%, 95% CI 2.1–7.3) The odds ratio of HIV infection in Table 1 also shows that age, marital status, household income, level of schooling, having a sexual partner outside of Likoma, and having a sexual partner older than 30 years were significantly associated with the probability of HIV infection in univariate analyses. Some relative

Table 1. Descriptive statistics for respondents of sexual network survey (*N* = 923).

	Women *N* = 501	Men *N* = 422	Odds ratio of HIV infection (95% CI)[b]
HIV test participation, HIV status and other health indicators			
Proportion of respondents who were			
Visited by HIV testing team	92.2%	93.3%	
Tested for HIV (among visited respondents)	73.9%	65.6%[a]	
HIV-positive (*N* = 597)	10.6%	4.7%[a]	
Proportion of respondents who			
Received an injection during the past year	55.0%	19.3%[a]	1.3 (0.67, 2.53)
Were ever tested for HIV before this study	22.2%	23.1%	0.97 (0.46, 2.02)
Reported symptoms of at least one STI	17.0%	17.3%	1.39[c] (0.69, 2.77)
Sociodemographic characteristics of respondents			
Age (years)			
< 20	22.5%	17.8%[a]	0.12 (0.01, 1.03)
20–24	31.6%	33.2%	1 (Ref.)
25–29	25.3%	21.1%	2.01 (0.87, 4.64)
30–34	13.0%	13.5%	2.13 (0.83, 5.45)
35+	7.6%	14.2%[a]	3.58 (1.34, 9.53)
Proportion born outside of Likoma	35.3%	31.7%	1.29[c] (0.7, 2.36)
Marital status			
Never married	36.9%	52.3%[a]	0.28 (0.12, 0.67)
Currently married	52.3%	44.9%[a]	1 (Ref.)
Widowed or divorced	10.8%	2.8%[a]	6.39 (2.83, 14.4)
Ever in a polygamous union[d]	10.5%	5.0%[a]	3.78 (1.45, 9.79)
Age at first marriage[d]			
≥ 20	36.6%	78.5%[a]	1 (Ref.)
< 20	63.4%	21.5%[a]	0.71[c] (0.36, 1.41)
Schooling			
Did not complete primary school	42.4%	46.5%	1 (Ref.)
Completed primary school	57.6%	53.5%	0.56 (0.31, 1.02)
Religion			
Anglican	82.8%	84.6%	1 (Ref.)
Other religions	17.2%	15.4%	1.29 (0.62, 2.68)
Household income			
< 10 000 MK/month (US$80)	71.0%	72.2%	1 (Ref.)
≥ 10 000 MK/month	29.0%	27.8%	1.8 (0.96, 3.36)
Sexual behaviors of respondents			
Number of sexual partners during 3-year period			
Before survey			
0–1	35.7%	26.1%[a]	1 (Ref.)
2	31.3%	25.4%[a]	1.2 (0.56, 2.58)
3 or more	33.0%	48.5%[a]	1.07 (0.5, 2.24)
Average difference between age of respondent and ages of partners			
Less than 5 years	57.2%	67.7%[a]	1 (Ref.)
More than 5 years	42.8%	32.3%[a]	1.3 (0.71, 2.37)
Proportion of respondents who report			
At least one sexual partner over 30 years old	45.8%	22.8%[a]	3.04 (1.61, 5.74)
At least one sexual partner outside of Likoma	30.7%	31.4%	1.9 (1.02, 3.5)
At least one new partner during last year	41.3%	59.1%[a]	0.7 (0.38, 1.29)
Having ever used a condom	68.5%	72.5%	0.86 (0.4, 1.85)
Using a condom consistently	17.4%	19.6%	0.45 (0.13, 1.54)

CI, Confidence interval; STI, sexually transmitted infection.
[a]The difference between women and men was significant at the 0.05 level, based on a chi-square test of association.
[b]Odds ratios of HIV infection were calculated using the Mantel–Haenszel formula for common odds ratios in stratified 2*2 tables, among the respondents tested for HIV (*N* = 597); the odds ratio reflect the prevalence of HIV relative to the reference category (for categorical variables), or relative to the zero ('no') value for binary variables.
[c]The null hypothesis of homogeneous odds ratios between men and women is rejected at the 0.05 level.
[d]Among ever-married respondents.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

risks differed significantly across sex [26]: for example, women who married relatively early were significantly less likely to be infected with HIV, and men presenting symptoms of STI or being born outside of Likoma were more likely to be infected.

Characteristics of reported relationships

Among the 1858 reports of sexual relationships, 1333 (72.7%) were to partners currently residing in Likoma, and 845 (45.5%) were to partners who were also interviewed during the ACASI survey (Table 2). The reliability of the reports of these 'within-sample' sexual relationships is relatively high, as 57.8% are reported jointly by both partners, with marital relationships being substantially more reliable than non-marital relationships (for further analyses of reporting patterns, see Helleringer *et al.* [26]). Based on these reports, we reconstructed a large sexual network including 1803 individuals: 923 respondents of the sexual network survey and 880 sexual partners of respondents who were not interviewed. These network members are connected by 1614 unique sexual relationships, as a total of 244 relationships among survey respondents were concordantly reported by both partners.

On average, male respondents had 2.6 sexual partners (IQR 1–3) during the 3 years before the survey, and female respondents had 2.2 partners (IQR 1–3; $P < 0.01$). Some 6.7% of women and 12.6% of men had five or more partnerships. Only very few respondents ($N = 27$) were not sexually active over the recall period. Marital relationships are stable, as 84.6% of marriages of women and 92.4% of marriages of men having started more than a year before the survey were still ongoing at the time of data collection. Virtually all marriages that started within a year of the survey were still ongoing. Non-marital relationships, on the other hand, are shorter. Only 40% of relationships that started during the past year were still ongoing at the time of the survey, and among non-marital relationships that started more than a year ago, only a quarter were still active. Only a small proportion of all non-marital relationships reported during the survey were characterized by respondents as one-off encounters.

Component size distribution

The sexual network identified by this study contains a total of 256 separate components. 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 that are shown in Fig. 2. Moreover, 883 network members, 56% of male and 45.6% of female survey respondents, constitute a single 'giant' component of individuals connected through sexual partnerships having taken place

Table 2. Descriptive statistics for sexual relationships reported by survey respondents ($N = 1858$).

	Women	Men
Among all relationship reports	$N = 884$	$N = 974$
Relationship is with a		
Marital partner	30.8% (272/884)	20.5%[a] (200/974)
Partner residing in Likoma	70.0% (617/884)	73.5%[a] (716/974)
Survey respondent ('within-sample partnership')	42.3% (374/884)	48.4%[a] (471/974)
Relationship is jointly reported by both partners (among within-sample partnerships only)	65.2% (244/374)	51.8%[a] (274/471)
Marital relationships	$N = 272$	$N = 200$
Marriage is among survey respondents ('within-sample')	60.6% (165/272)	80.0%[a] (160/200)
Jointly reported by both spouses	94.0% (155/165)	96.9 (155/160)
Marriage is ongoing at survey, among all marital relationships that started		
Within one year before survey	100.0% (22/22)	96.3% (26/27)
More than one year before survey	84.6% (208/246)	92.4%[a] (158/171)
Non-marital sexual relationships	$N = 612$	$N = 774$
Relationships is among survey respondents ('within-sample')	34.2% (209/612)	40.2%[a] (311/774)
Jointly reported by both partners	42.6% (89/209)	28.6%[a] (89/311)
Relationship is with		
Steady boyfriend or girlfriend	57.8% (344/595)	54.5% (401/736)
Infrequent partner	33.6% (200/595)	33.5% (247/736)
One-off encounters	8.6% (51/595)	12.0% (88/736)
Starting date of extramarital relationships		
Within past month	7.6% (46/605)	6.6% (51/767)
Within past year	26.8% (162/605)	29.1% (223/767)
More than a year ago	65.6% (397/605)	64.3% (493/767)
Relationship is ongoing at survey, among all non-marital relationships that started		
Within one year before survey	43.9% (89/203)	39.7%[b] (108/272)
More than one year before survey	27.4% (107/391)	20.7%[b] (102/488)

The numerators and denominators for the calculation of the percentage are given in parenthesis. Discrepancies in sample sizes are the result of missing data, as some respondents failed to answer certain questions and are thus not included in the counts of individuals falling into particular categories.
[a]The difference between women and men was significant at the 0.05 level, based on a chi-square test of association.
[b]The difference between women and men was significant at the 0.1 level, based on a chi-square test of association.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

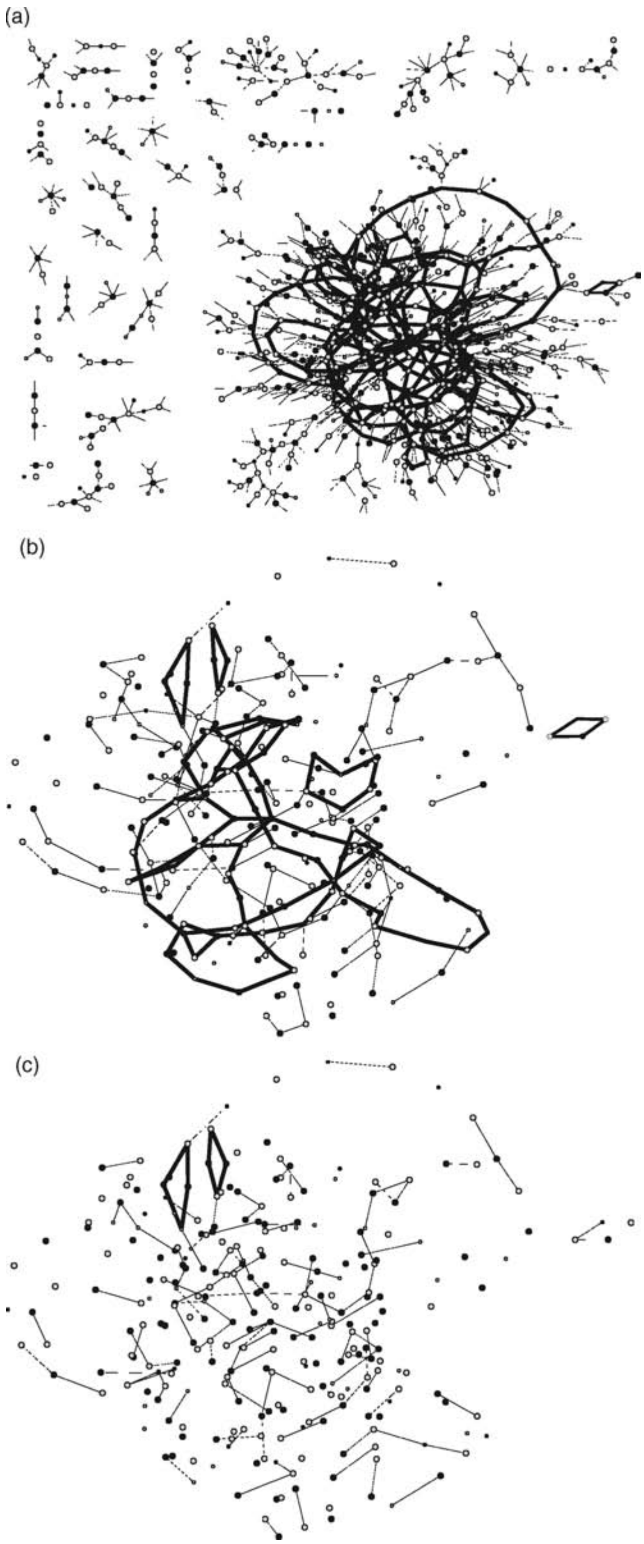

Fig. 2. Components of the Likoma sexual networks of size six and larger. Circles represent individuals. Lines represent sexual partnerships between individuals. Black circles: male survey respondents; grey circles: female survey respondents. Larger circles represent network members who were interviewed during the sexual network survey and who were sexually active during the recall period ($n = 896$). Smaller circles represent

during the 3 years before the survey. Inclusion in this component is not necessarily associated with a large numbers of partners: whereas members of the giant component on average reported a higher number of partners than other respondents (3 versus 1.8; $P < 0.01$), a substantial fraction of members of the giant component (40%) also had at most two partners during the 3 years before the survey. The connectivity of the sexual network therefore occurs not because a small number of individuals have high rates of partner change [35], 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.

High prevalence of cycles in the network

The giant component dominating the sexual network among young adults in the study villages (Fig. 2a) also contains three bicomponents that include individuals who are connected by more than one independent path. Two of these bicomponents are simple 'diamonds' connecting four network members in a cycle. The third bicomponent, however, is significantly larger and includes 274 individuals, that is, 15% of all members of the sexual network and close to 25% of survey respondents. Among respondents below the age of 25 years, 36% belong to this denser region of the network. Several young adults in our study villages have had more than one partner in common, and a significant proportion of inhabitants of Likoma may be at an increased risk of HIV infection because they are connected by multiple independent chains of sexual relationships. Moreover, the occurrence of cyclical structures in the observed sexual network does not depend on the relatively long recall period of 3 years: the bicomponents in Fig. 2 contain cycles even if only recent or currently active relationships

Fig. 2 *(continued)*
network members who were found within the village rosters but were not interviewed because they were outside the sampling frame of this study (all young adults aged 18–35 years and their spouses living in the seven sample villages shown in Fig. 1). The subset of lines not connecting two circles represent partnerships with individuals we were not able to identify in the rosters of potential partners. The subset of thicker lines represent partnerships within bicomponents, that is, between network members who are connected by more than one independent pathway within the sexual network. (a) Full network of relationships identified during this study (3-year recall period). The right-most panels describe the evolution of the network bicomponent over time. (a) All sexual relationships in components of size 6 or larger that were ever active during the 3 years before the survey. Relationships within the bicomponents of (a) that were active. (b) Within 1 year before the survey. (c) At the time of the survey. Subsets of thicker relationships in these graphs represent bicomponents within the time interval under consideration. The location of circles in this graph is not related to the geographical location of respondents within the island.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

are considered. For example, the network of relationships that were active within one year of the survey contains several bicomponents connecting a total of 84 network members (Fig. 2b). Short-length cycles (i.e. two individuals having two partners in common) are also present within the network of relationships that were ongoing at the time of the survey (Fig. 2c). For comparison, Bearman *et al.* [25] did not find any short-length cycles among students of a US high school over an 18 month period.

Network location and HIV infection

Analyses of the position of HIV-positive individuals within the network indicate that the distribution of HIV is not homogeneous (Table 3), ranging from 3% in the bicomponents to 8.9% in the disjoint components and 10.8% in the 'branches' of the giant component. Whereas members of the bicomponents (Fig. 2) are exposed to potentially multiple pathways of HIV infection, HIV prevalence is highest in the sparser regions of the network, that is, small components and the giant component outside the bicomponents. Several factors contribute to this apparently paradoxical distribution of HIV prevalence. First, the sociodemographic composition of the sparser regions of the network favors an increased prevalence of HIV as some groups who are more likely to be infected, for example, older respondents, women and widows (Table 1), are overrepresented in these regions. Second, the prevalence of several risk factors associated with HIV infection also varied significantly within the network (Table 3). For example, the proportion of respondents having engaged in relationships outside of Likoma, or having engaged in a relationship with a partner older than 30 years, were

Table 3. Prevalence of HIV and risk factors for HIV infection across network locations.

	Network location		
	Small components	Main component	Bicomponents
Mean age	27.5	23.4	22.1
(IQR)	(23, 32)	(19, 27)	(19, 25)
Proportion of network members who are			
Female	58.9	52.4	46.9[a]
Never married	29.1	55.5	63.5[a]
Divorced or widowed	10.6	3.9	3.15[a]
Average number of partners			
Men	2.1	2.3	3.8[a]
(SD)	(1.6)	(1.2)	(1.4)
Women	1.4	2.2	3.7[a]
(SD)	(0.8)	(1.1)	(1.3)
Health variables[b]			
Proportion of respondents who			
Were tested for HIV	65.8	72.3	77.0[a]
Were infected with HIV	8.9 (5.9, 11.8)	10.8 (6.3, 15.3)	3.0 (0.6, 5.9)
Reported symptoms of STI	16.7 (13.1, 20.4)	16.5 (12.1, 20.8)	17.5 (12.3, 22.6)
Received an injections within one year before survey	34.2 (29.9, 38.5)	40.1 (34.4, 45.8)	35.0 (29.3, 40.7)
Were ever tested for HIV before this survey	21.9 (18.0, 25.9)	21.1 (15.9, 26.3)	25.3 (19.7, 30.8)
Sexual mixing variables[c]			
Proportion of respondents reporting			
Any partner outside of Likoma	40.7 (36.5, 44.9)	25.3 (20.6, 30.0)	16.9 (14.5, 19.3)
Any partner above 30 years old	35.8 (32.6, 38.9)	23.8 (20.0, 27.6)	12.6 (10.7, 14.6)
Consistent condom use	16.7 (12.9, 20.6)	16.7 (11.9, 21.5)	11.9 (8.8, 15.0)
Proportion of relationships that are ongoing at survey among all			
Marriages that started more than one year before survey	87.2 (83.2, 91.2)	93.5 (88.6, 98.4)	68.9 (60.2, 77.5)
Extramarital relations that started more than one year before survey	28.1 (22.0, 34.0)	25.6 (20.2, 31.0)	21.6 (17.3, 26.0)
Extramarital relations that started within one year before survey	43.7 (35.1, 52.2)	42.7 (34.8, 50.6)	40.9 (35.3, 46.4)

STI, Sexually transmitted infection. Network locations: small components: members of small components (isolates, dyads, triads, etc.); main component: members of the giant component, outside the bicomponents (Fig. 2a); bicomponents members of the bicomponents within the giant component (individuals connected by thick lines in Fig. 2). 95% confidence intervals are in parentheses unless otherwise noted.
[a]The difference in proportions between network locations is significant at the 0.05 level.
[b]Proportions are standardized by age, sex and marital status.
[c]Proportions are standardized by age, sex, marital status and number of partners.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

higher in the sparser regions of the network. Relationships, in particular marital relationships, tend to be longer in the smaller components, and as a result concurrent partnerships might actually be more common in these structures. Table 3 also shows that the participation in HIV testing differed across regions of the network, but this modest variation is unlikely to explain the differential HIV prevalence across regions of the network.

Discussion

This study constitutes, to the best of our knowledge, the first sociocentric study of sexual networks among a general population of sub-Saharan Africa. Combining a census of the population of Likoma Island on Lake Malawi with a saturated sexual network survey of all inhabitants aged 18–35 years in geographically contiguous villages on the island, we reconstructed the quasi-complete sexual networks of young adults in the sample villages. Two main findings emerge from our analyses. First, we document the existence of a large and robust sexual network connecting a substantial fraction of the island's young adult population. Half of all sexually active respondents were linked together in a giant network component, and more than a quarter were members of bicomponents. They were thus connected together through multiple independent chains of sexual relationships. Such structural features of sexual networks have been associated with the spread of STI in theoretical investigations [36], but before this study, these network structures had only been identified in analyses of STI transmission among core group members [23,37]. A key finding of our analyses is that these structures also exist within the general population of Likoma, that is, a population in which reported rates of partner change are low (Table 1), and heterogeneity in numbers of partners among population members is limited.

The high connectivity of this sexual network is also not merely the result of the long recall period of 3 years used during the study: robust cycles also exist within the network of recent and currently active sexual relationships. Contrary to claims that sexual networks outside of core groups are too sparse to sustain generalized epidemics through heterosexual contacts in remote areas of sub-Saharan Africa [38,39], the sexual network structure observed on Likoma Island appears to be compatible with a broad diffusion of pathogens within the population of the island, independently of parenteral transmission of HIV during healthcare.

Second, our analyses reveal important differences in the structural position of HIV-positive individuals. The prevalence of HIV varied significantly across locations of the network, with sparser regions having a higher prevalence than densely connected bicomponents. Several factors contribute to this ecological variation of HIV prevalence within the network. For example, the sparser regions of the sexual network are characterized by an overrepresentation of socioeconomic groups with increased prevalence (e.g. older respondents, women, widows), and a greater exposure to external infections through older partners or partners from the mainland.

There are several other processes that could help explain the distribution of HIV within the network. Individuals may get infected when they are closer to the dense regions of the networks but subsequently 'drift' into smaller disjoint components. Such trajectories may occur because declining health at advanced stages of the disease limits an individual's physical ability to engage in sexual relationships, or because HIV-positive individuals may have fewer opportunities for sexual relationships (stigma, partner selection). On the other hand, relationships are possibly longer in smaller components (Table 2), and as a result the concurrency of partnerships might be more widespread in these structures. Men in the small components also have significantly more partnerships than women (Table 3), therefore leading to starlike network patterns with many low activity individuals attached to one more active individual. As suggested by several mathematical models [16,40], increased behavioral heterogeneity together with longer partnership durations could thus help explain the higher rates of HIV infection observed in the smaller components of the network, particularly among women. Longitudinal data as well as more precise measures of partnership timing and duration would, however, be required to confirm the importance of these dynamic processes on the incidence of HIV within populations.

Although only sociocentric network studies can identify the population-level structural characteristics of sexual networks and the position of HIV-positive persons within these networks, several limitations of our study are important when interpreting the above findings. First, despite its epidemiological relevance, the recall period of 3 years is prone to forgetting of partners [41]. Incomplete network bias also occurs because at most five partners were recorded in the survey, and partners older than 35 years of age (except spouses) or living outside the sample villages were not interviewed at all. As a result, our study underestimates the connectivity of the sexual networks in which young adults in Likoma are embedded. Our estimates of the size of the giant component and the bicomponents in Fig. 2 are thus conservative, and some of the smaller or sparser network components may actually be attached to larger or denser network structures [24]. For example, 35% of dyads, 73% of triads, and 80% of components of size four include a network member who was not interviewed. Finally, the island context of this study potentially limits our ability to generalize findings to a broader array of contexts in sub-Saharan Africa. In regions of the mainland where transportation is more readily available or that are located closer to major cities, the pool

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

of potential sexual partners might be less limited than in Likoma. The cyclical structures we identified might thus not emerge on the same scale or within the same time frames in different socioeconomic contexts. Further research using longitudinal and comparative research designs is thus needed properly to assess the role of sexual network structures in the diffusion and prevention of HIV in sub-Saharan Africa.

Acknowledgements

The authors would like to thank Agnes Chimbiri, Paul Hewett, Agatha Bula and George Joaki for their contributions to this project. They also thank the inhabitants of Likoma for their collaboration in this study.

Sponsorship: This research received support through the National Institute of Child Health and Development (grant nos. RO1 HD044228 and RO1 HD/ MH41713), National Institute on Aging (grant no. P30 AG12836), the Boettner Center for Pensions and Retirement Security at the University of Pennsylvania, and the National Institute of Child Health and Development Population Research Infrastructure Program (grant no. R24 HD-044964), all at the University of Pennsylvania, and the Norwegian Center for Advanced Studies.

Conflicts of interest: None.

References

1. Weir S, Pailman C, Mahlalela X, Coetzee N, Meidany F, Boerma J. **From people to places: focusing AIDS prevention efforts where it matters most.** *AIDS* 2003; **17**:895–903.
2. Glynn JR, Ponnighaus J, Crampin AC, Sibande F, Sichali L, Nkhosa P, *et al*. **The development of the HIV epidemic in Karonga district, Malawi.** *AIDS* 2001; **15**:2025–2029.
3. Morris CN, Ferguson AG. **Estimation of the sexual transmission of HIV in Kenya and Uganda on the Trans-Africa Highway: the continuing role for prevention in high risk groups.** *Sex Transm Infect* 2006; **82**:368–371.
4. Carael M, Slaymaker E, Lyerla R, Sarkar S. **Clients of sex workers in different regions of the world: hard to count.** *Sex Transm Infect* 2006; **82 (Suppl 3)**:iii26–iii33.
5. Gomes do Espirito Santo M, Etheredge GD. **Male clients of brothel prostitutes as a bridge for HIV infection between high risk and low risk groups of women in Senegal.** *Sex Transm Infect* 2005; **81**:342–344.
6. Gouws E, White PJ, Stover J, Brown T. **Short term estimates of adult HIV incidence by mode of transmission: Kenya and Thailand as examples.** *Sex Transm Infect* 2005; **82 (Suppl 3)**:51–55.
7. Wasserheit JN, Aral SO. **The dynamic topology of sexually transmitted disease epidemics: implications for prevention strategies.** *J Infect Dis* 1996; **174 (Suppl 2)**:201–213.
8. Blanchard JF. **Populations, pathogens, and epidemic phases: closing the gap between theory and practice in the prevention of sexually transmitted diseases.** *Sex Transm Infect* 2002; **78 (Suppl 1)**:183–188.
9. Aral SO, Blanchard JF. **Phase specific approaches to the epidemiology and prevention of sexually transmitted disease.** *Sex Transm Infect* 2002; **78 (Suppl 1)**:1–2.
10. Doherty IA, Padian NS, Marlow C, Aral SO. **Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections.** *J Infect Dis* 2005; **191**:S42–S54.
11. Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T, Caraël M, *et al*. **Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe.** *Lancet* 2002; **359**:1896–1903.
12. Kelly R, Gray R, Sewankambo N, Serwadda D, Wabwire-Mangen F, Lutalo T, *et al*. **Age differences in sexual partners and risk of HIV-1 infection in rural Uganda.** *J Acquir Immune Defic Syndr* 2003; **32**:446–451.
13. Lagarde E, Schim van der Loeff M, Enel C, Holmgren B, Dray-Spira R, Pison G, *et al*. **Mobility and the spread of human immunodeficiency virus into rural areas of West Africa.** *Int J Epidemiol* 2003; **32**:744–752.
14. Coffee M, Garnett G, Mlilo M, Voeten H, Chandiwana S, Gregson S. **Patterns of movement and risk of HIV infection in rural Zimbabwe.** *J Infect Dis* 2005; **191 (Suppl 1)**:S159–S167.
15. Kishamawe C, Vissers DC, Urassa M, Isingo R, Mwaluko G, Borsboom GJ, *et al*. **Mobility and HIV in Tanzanian couples: both mobile persons and their partners show increased risk.** *AIDS* 2006; **20**:601–608.
16. Morris M, Kretzschmar M. **Concurrent partnerships and the spread of HIV.** *AIDS* 1997; **11**:641–648.
17. Lagarde E, Auvert B, Caraël M, Laourou M, Ferry B, Akam E, *et al*. **Concurrent sexual partnerships and HIV prevalence in five urban communities of sub-Saharan Africa.** *AIDS* 2001; **15**:877–884.
18. Schneeberger A, Mercer C, Gregson S, Ferguson N, Nyamukapa C, Anderson R, *et al*. **Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe.** *Sex Transm Dis* 2004; **31**:380–387.
19. Jones JH, Handcock MS. **Sexual contacts and epidemic thresholds.** *Nature* 2003; **423**:605–606.
20. Handcock MS, Jones JH. **Likelihood-based inference for stochastic models of sexual network formation.** *Theoret Popul Biol* 2004; **65**:413–422.
21. Latora V, Nyamba A, Simpore J, Sylvette B, Diane S, Sylvere B, *et al*. **Network of sexual contacts and sexually transmitted HIV infection in Burkina Faso.** *J Med Virol* 2006; **78**:724–729.
22. Ghani A, Ison C, Ward H, Garnett G, Bell G, Kinghorn G, *et al*. **Sexual partner networks in the transmission of sexually transmitted diseases: an analysis of gonorrhea cases in Sheffield, UK.** *Sex Transm Dis* 1996; **23**:498–503.
23. Rothenberg RB, Potterat JJ, Woodhouse DE, Muth SQ, Darrow WW, Klovdahl AS. **Social network dynamics and HIV transmission.** *AIDS* 1998; **12**:1529–1536.
24. Wylie J, Cabral T, Jolly A. **Identification of networks of sexually transmitted infection: a molecular, geographic, and social network analysis.** *J Infect Dis* 2005; **191**:899–906.
25. Bearman PS, Moody J, Stovel K. **Chains of affection: the structure of adolescent romantic and sexual networks.** *Am J Sociol* 2004; **110**:44–91.
26. Helleringer S, Kohler HP, Chimbiri A, Chatonda P, Mkandawire J. The Likoma Network Study: context, data collection and initial results. PSC Working Paper Series PSC 07-05. 2007; University of Pennsylvania, Population Studies Center. Available at: http://repository.upenn.edu/psc_working_papers/9. Accessed: July 2007
27. Ghani AC, Donnelly CA, Garnett GP. **Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases.** *Stat Med* 1998; **17**:2079–2097.
28. Bloom DE. **Technology, experimentation, and the quality of survey data.** *Science* 1998; **280**:847–848.
29. Mensch BS, Hewett PC, Erulkar A. **The reporting of sensitive behavior among adolescents: a methodological experiment in Kenya.** *Demography* 2003; **40**:247–268.
30. Heuveline P. **HIV and population dynamics: a general model and maximum-likelihood standards for East Africa.** *Demography* 2003; **40**:217–245.
31. Blasnick M. Gsoundex: Stata module to implement soundex algorithm. 2001; Boston College Department of Economics. Available at: http://ideas.repec.org/c/boc/bocode/s420901.html. Accessed: June 2007
32. Wasserman S, Faust K. *Social network analysis: methods and applications*. Cambridge: Cambridge University Press; 1994.
33. De Nooy W, Mrvar A, Batagelj V. *Exploratory social network analysis with Pajek*. Cambridge: Cambridge University Press; 2005.
34. Moody J. **The importance of relationship timing for diffusion.** *Social Forces* 2002; **81**:25–46.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

35. Liljeros F, Edling CR, Amaral LAN, Stanley HE, Aberg Y. **The web of human sexual contacts.** *Nature* 2001; **411**:907–908.
36. Newman M. **Spread of epidemic disease on networks.** *Phys Rev E* 2002; **66**:16128–16211.
37. Potterat JJ, Phillips-Plummer L, Muth SQ, Rothenberg RB, Woodhouse DE, Maldonado-Long TS, *et al*. **Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs.** *Sex Transm Infect* 2002; **78 (Suppl 1)**:159–163.
38. Hudson CP. **AIDS in rural Africa: a paradigm for HIV-1 prevention.** *Int J STD AIDS* 1996; **7**:236–243.
39. Brewer DD, Brody S, Drucker E, Gisselquist D, Minkin SF, Potterat T, *et al*. **Mounting anomalies in the epidemiology of HIV in Africa: cry the beloved paradigm.** *Int J STD AIDS* 2003; **14**:144–147.
40. Morris M, Kretzschmar M. **A microsimulation study of the effect of concurrent partnerships on the spread of HIV in Uganda.** *Math Popul Stud* 2000; **8**:109–133.
41. Brewer DD, Webster CM. **Forgetting of friends and its effects on measuring friendship networks.** *Soc Netw* 1999; **21**:361–373.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.