Sexual Networks, Surveillance, and Geographical Space during Syphilis Outbreaks in Rural North Carolina

Irene A Doherty\textsuperscript{a}, Marc L Serre\textsuperscript{b}, Dionne Gesink\textsuperscript{c}, Adaora A Adimora\textsuperscript{a,b}, Stephen Q Muth\textsuperscript{d}, Peter A Leone\textsuperscript{a}, and William C Miller\textsuperscript{a,b}

\textsuperscript{a}School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill NC

\textsuperscript{b}Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill NC

\textsuperscript{c}Dalla Lana School of Public Health, University of Toronto, Toronto Ontario, Canada

\textsuperscript{d}Quintus-ential Solutions, Colorado Springs, CO

Abstract

\textbf{Background}—Sexually transmitted infections (STIs) spread along sexual networks whose structural characteristics promote transmission that routine surveillance may not capture. Cases who have partners from multiple localities may operate as spatial network bridges, thereby facilitating geographical dissemination. We investigated the relationships between surveillance, sexual networks, and spatial bridges for syphilis outbreaks in rural counties of North Carolina.

\textbf{Methods}—We selected from the state health department’s surveillance database cases diagnosed with primary, secondary, or early latent syphilis during October 1998 - December 2002 residing in central and southeastern North Carolina, along with their sex partners, and their social contacts irrespective of infection status. We applied matching algorithms to eliminate duplicate names and create a unique roster of partnerships from which networks were compiled and graphed. Network members were differentiated by disease status and county of residence.

\textbf{Results}—In the county most affected by the outbreak, densely connected networks indicative of STI outbreaks were consistent with elevated incidence and a large case load. In other counties, the case loads were low with fluctuating incidence, but network structures suggested the presence of outbreaks. In a county with stable, low incidence and a high number of cases, the networks were sparse and dendritic, indicative of endemic spread. Outbreak counties exhibited densely connected networks within well-defined geographic boundaries and low connectivity between counties; spatial bridges did not seem to facilitate transmission.

\textbf{Conclusions}—Simple visualization of sexual networks can provide key information to identify communities most in need of resources for outbreak investigation and disease control.

Public health practitioners often rely on epidemiologic measures of disease occurrence, including the absolute number of cases, prevalence, and incidence, to detect local epidemics of sexually transmitted infections (STIs). A moderately high incidence in a densely
populated urban community likely involves a large number of cases. In less populated rural areas, however, surveillance statistics alone may not accurately distinguish between an outbreak and a random fluctuation in the case number. When a relatively small number of people in a rural area become infected or the number of new cases increases slightly, incidence may rise alarmingly, incorrectly signaling the emergence of an outbreak. Alternatively, a stable low incidence does not necessarily rule out the possibility of an outbreak among a small high-risk segment of the population within the catchment area.

STIs are transmitted along sexual networks. Structural characteristics of networks help determine transmission patterns, which are not always identifiable with traditional epidemiologic measures. During an outbreak, densely connected networks composed of cases in several short-term or concurrent partnerships create network microstructures that support rapid transmission. In contrast, populations in an endemic state typically have networks with a predominance of dyads, triads, and other small disconnected dendritic or radial structures. Bridging is another characteristic of sexual networks that affects the reach of transmission. Potential examples of sexual bridging include bisexual men, injection-drug users, and people involved with transactional sex. Concurrent sexual partnerships are a form of temporal bridging. Spatial bridges arise among infected people who have partners in more than one neighborhood, county, or state, thereby facilitating geographical spread.

We have described the sexual and social network qualities of a heterosexually-transmitted outbreak of syphilis in a largely rural region of southeastern North Carolina that spanned several counties. Previous analyses focused on characteristics and behaviors of individuals, and sexual partnership patterns and dynamics. Factors that fostered the growth of the epidemic include widespread crack cocaine use; exchange of sex for drugs and money; and sexual mixing with respect to age and race/ethnicity, number of sex partners, and the stage of the syphilis diagnosis (i.e., primary, secondary or early latent).

Here, we extend our investigation to the structural features of the networks as they relate to surveillance and geographical space. First, the analyses will illustrate how sexual network assessment augments use of traditional epidemiologic measures of disease occurrence to identify where outbreaks of STIs are underway. Second, we will show that spatially-localized outbreaks occurred within well-defined geographic boundaries and that bridges between partners residing in different counties were insufficient to account for the rapid growth of syphilis across communities and the region.

**METHODS**

**Overview**

Traditional surveillance measures include county-level annual incidence and number of newly diagnosed cases per year by county. Network components are subsets of connected individuals (nodes) within the broader sexual-social network, thus a collection of components comprise pieces of one network. The distribution of sizes of components is an informative assessment of structural features of the overall network. We produce graphs of the network components to depict the constellation of sexual and social ties among nodes.
With inclusion of appropriate data, nodes are distinguished by disease status (infection, confirmed uninfected, and unknown) and by county of residence.

**Surveillance measures**

All newly diagnosed cases of syphilis infection are reported to the county and state health departments for surveillance and implementation of disease control protocols. We present the annual incidence and absolute number of cases reported in publications produced by the state STD/HIV Prevention and Treatment Branch of the Division of Public Health during 1998-2002 for the following North Carolina counties: Moore, Montgomery, Robeson, Columbus, and Cumberland (Figure 1).

**Network assessment**

Disease Intervention Specialists interview each case to elicit information about sex partners and social contacts at risk of acquiring or transmitting infection. These staff persons also perform contact tracing, notify partners of potential exposure, and offer testing and prophylactic treatment. The Centers for Disease Control and Prevention recommends that they perform “cluster interviews” with groups of infected and uninfected people affiliated with each another (either as social or sex partners) to obtain names of anyone associated with cases who may have symptoms or be at risk of becoming infected. These interviews can sometimes lead to identification of additional sex partners who were not initially revealed by cases during their first conversation. Data for each person named (including anonymous partners), regardless of infection status, and each sexual partnership and social tie is entered into an electronic database, thereby permitting construction and assessment of social and sexual networks. Common names, nicknames, and individuals named multiple times can distort networks; one of the authors (SQM) developed algorithms to eliminate duplicate names and produce a file of unique persons and a roster of social and sexual ties.

Available data include cases diagnosed with primary, secondary, or early latent syphilis during October 1998 - December 2002 residing in central and southeastern NC, and cases’ sex partners and social contacts irrespective of infection status. Cases younger than 14 years of age (which typically represent newborns diagnosed with congenital syphilis) and patients with diagnoses of late latent syphilis (indicative of prolonged subclinical infection) were excluded. We classified people as uninfected if the database documented negative syphilis test results or administration of prophylactic treatment. Sex partners and social contacts who either refused testing or whom the Disease Intervention Specialists could not locate were classified as having unknown disease status.

We used UCINET and NetDraw to compile and graph the structure of the network. Visual inspection of the diagrams permitted us to distinguish between network (and component) structures consistent with outbreaks or representative of endemicity. Then we examined the county of residence for each node in the components to determine whether the components were geographically localized.
RESULTS

Epidemiologic surveillance

The statewide incidence of early-stage syphilis, which includes primary, secondary, and early latent syphilis, fell from 20 cases per 100,000 person-years in 1998 to 7.4/100,000 person-years in 2002 (Figure 2A). During the same period, the incidence of early-stage syphilis in Robeson County, the apparent epicenter of the outbreak, increased from 87/100,000 person-years to a high of 116/100,000 person-years by 2001, with between 105 and 144 new cases per year. Columbus County, which borders Robeson County (Figure 1), experienced a 600% increase as incidence climbed from 15/100,000 person-years (1998) to a peak of 98.5/100,000 person-years in 2001. After several years of incidence less than 12/100,000 person-years and fewer than 15 cases per year in Montgomery County, the 52 cases diagnosed during 2000 corresponded to an incidence of 194/100,000 person-years. During 2001-2002, syphilis incidence in Moore County, which borders Montgomery County (Figure 1), doubled. The absolute number of diagnoses in Cumberland County exceeded all of the other counties except Robeson County; however, incidence remained at low, endemic levels (Figure 2B).

Cases and contacts

Of the 5,299 persons in the database, 2055 (39%) were infected cases, 2497 (47%) were confirmed as uninfected or received prophylactic treatment, and 747 (14%) had unknown infection status (Table 1). The percent who were infected was higher for women (43%) than men (36%). However, the status of 16% of men was not determined, as compared with 12% of women. Disease Intervention Specialists traced proportionately more contacts in the epicenter, Robeson County, and its neighbor, Columbus County; about 10% of contacts in these counties were not traced or tested, in contrast to 16% in the other counties. The median age was 35 years for men and 30 years for women. In a previous analysis, we determined that the average age difference between men and their female partners was 6.6 years.

Network and components

The network comprised 4,833 sexual and social partnerships involving at least one infected person, distributed across 907 separate components (Table 2). Dyads composed the largest proportion of components (43%) across both outbreak and non-outbreak counties, yet only 15% of people had a single partner (i.e., were in a dyad). Similarly, 24% of the components were triads, distributed among 12% of people. The two largest components involved 895 and 278 nodes, the first accounting for 17% of all persons and the second accounting for 5%.

The component with 278 nodes, in Columbus County (Figure 3), included 68 cases, 182 uninfected people, and 28 persons with unknown infection status. This component was highly geographically localized; of the total 314 county residents in the database, 79% (regardless of infection status) were part of this component. Furthermore this single component included 67% of the 99 cases from Columbus County. Although Columbus County borders Robeson County (considered the epicenter of the outbreak), between-county
connections were virtually absent – only two cases and seven uninfected partners from Robeson County were part of this component.

The largest component included 895 people – 310 cases (35%), 474 uninfected persons (53%), and 111 (12%) people with unknown disease status. This component has two large substructures (Figures 4A and 4B). Most of the network members resided in three counties: Robeson (66%), Montgomery (17%), or Moore (10%). Montgomery and Moore County border each other and Richmond County is located between these two counties and Robeson County (Figure 1). The person who connected (or bridged) the two substructures resided in Richmond County. This node and its contacts are shown in both figures with an arrow to indicate where the linkage appears in the component. In preliminary temporal analysis (not shown), we determined that this person was diagnosed about halfway through the observation period. Networks in the surrounding counties had already grown into outbreaks.

The size and structure of the sexual network in Robeson County (Figure 4A) clearly depicts a massive outbreak, as confirmed by the incidence (Figure 2A) and the absolute number of cases (Figure 2B).

Montgomery County experienced a sudden increase in incidence in 2000 associated with a relatively small number of cases (Figure 2), which may have been a random fluctuation. However, assessment of the associated network revealed interconnections and cohesiveness, indicative of an epidemic phase (Figure 4B).\textsuperscript{1,9,10} Compared with the other three counties (Robeson [Figure 4A], Montgomery [Figure 4B], and Columbus [Figure 3]), cases in Moore County formed relatively fewer closed loops, and incidence did not increase until the last year of observation (2001-2002). Nonetheless, structural characteristics of connections among Moore County residents suggest the presence of an outbreak. Furthermore, few partnerships connected Moore and Montgomery County.

In Cumberland County, which also borders Robeson County, neither the incidence nor the structures of the sexual networks were consistent with an epidemic (Figure 5), despite the large number of cases (Figure 2B). Dyads comprised \textit{39\%} of all network components from Cumberland County; \textit{23\%} of residents in our data had one partner. The largest component included 29 people in a dendritic form (Figure 5). Other components were radial, and only two components involved cases in closed cycles. The network structures in this collection of disconnected components indicate an endemic state; preliminary temporal analyses showed that these components persisted throughout the observation period.

**DISCUSSION**

Sexual network assessment can identify STI outbreaks by simply displaying transmission pathways that traditional epidemiologic measures of disease occurrence do not necessarily capture. In a large predominately rural region of North Carolina, an epidemic of heterosexually transmitted syphilis actually consisted of several smaller outbreaks characterized by high network interconnectivity within each county, and low connectivity between counties, resulting in spatially-localized outbreaks within well-defined geographic boundaries.
Localized prevalence and incidence provides an incomplete picture of STI epidemiology because the patterns of sexual partnerships and structures of sexual networks are hidden. A new picture emerges when we combine incidence and case-load data with sexual-network data. The high incidence and concomitant high case number support the occurrence of an outbreak in Robeson County, which the sexual network graph confirms (Figure 4). However, outbreaks in Montgomery and Columbus counties in 2000 and 2001 are less clearly delineated. From the transient high incidence yet low number of cases, we cannot be certain without examining the sexual networks (Figure 4A), whether the fluctuations were random or the result of active outbreaks. The partnership patterns in the component for these two counties were sufficiently cohesive to foster epidemic growth. In fact, at the time of the regional outbreak, Montgomery County was not recognized as being involved (PA Leone, 2 May 2011, personal communication). Conversely, in Cumberland County, where the case load was high and the incidence relatively low, dyads and small branching components predominated (Figure 5), suggesting that the county was in an endemic phase and no subgroup of the county’s residents experienced an outbreak.

As acknowledged in our previous analysis of these data, one of the key limitations of this analysis is incomplete network ascertainment because: (1) some cases may not have named partners; (2) Disease Intervention Specialists were not given enough information to locate some named partners; (3) named partners may have declined to talk to Disease Intervention Specialists; (4) named partners living in another state could not be traced; and (5) named contacts’ identities may not have been correctly matched to those of other people in the network, effectively concealing the presence of ties between people. Despite the multiple sources of missing data, network structures in this epidemic are consistent with past network studies involving STI outbreaks. The missing partnerships in any of the outbreak counties would not dissuade decision makers from deploying additional Disease Intervention Specialists to these communities. In Cumberland County, where fragmented small components predominated, the addition of missing nodes and partnerships would not substantially alter their structure.

The data are retrospective and the network components appear as cross-sectional in this analysis, even though the cross-sectional data actually represent the cumulative growth of the outbreaks. Viewing temporal growth of the network, as new cases become diagnosed and enter the network, would produce a more complete picture of the outbreaks. But this finer view comes at the cost of completeness and could have limited the ability to discern certain patterns within the network that revealed where active transmission took place.

Using a combination of classical surveillance and network examination could inform public-health decision-makers where to direct (what are often limited) disease control resources. Implementing the existing evidenced-based programs that prevent transmission is critically important for syphilis. After decades of near-elimination in much of the United States, syphilis re-emerged during 2000-2001 among men who have sex with men, and syphilis incidence continues to escalate. Co-epidemics of syphilis and HIV among men who have sex with men exacerbate transmission of both pathogens. Moreover, the outbreaks in Robeson, Columbus, and nearby counties were tied to transactional sex, crack cocaine use,
and heterosexual transmission, as have been other outbreaks in the southeast.\textsuperscript{45} A similar heterosexual outbreak in another county in NC occurred during 2008-2010.\textsuperscript{46} Thus, although we describe an outbreak from 10 years ago, the findings have direct relevance to public-health practice today.

Disease Intervention Specialists meet with each new patient diagnosed with syphilis, elicit the names of the patients’ sex partners, perform contact tracing, and interview and test each partner. This process is repeated for each new person diagnosed during the course of investigations. During outbreaks, Disease Intervention Specialists may perform additional cluster interviews with infected and uninfected persons who are socially or sexually connected to cases, and those interviews may lead to the identification of additional cases. These tasks are labor-intensive and costly. If public-health practitioners could recognize the locales where sexual-network structures are developing in a manner that is indicative of outbreaks, they could prioritize where to place resources, regardless of the relative number of new diagnoses.

Simple visualization of networks permits public-health professionals to better understand the underlying epidemiology of STIs. Mapping sexual networks in social space and geographically could help Disease Intervention Specialists connect cases to each other. Although this approach currently would be burdensome and add to the work of Disease Intervention Specialists, our results underscore the need to develop and improve on-the-ground tools for STI control. While some user-friendly and inexpensive (or free) software is increasingly available (such as NodeXL http://nodexl.codeplex.com\textsuperscript{)}, adapting it for STI control use is not straightforward. Nevertheless, with the proliferation of “apps” for smart phones to find nearby restaurants, parks, hotels, or even potential sex partners, we believe similar technology can and should be developed specifically for public-health departments and use by Disease Intervention Specialists. It would provide an important adjunct to the conventional measures of epidemiologic surveillance. The ability to assess networks in near or real time may help guide how best to use disease control resources.

\textbf{Acknowledgments}

We thank the HIV/STD Prevention & Treatment Branch of the North Carolina Division of Public Health for granting us access to these data.

\textbf{Sources of financial support:} Supported by a grant from the National Institutes of Health grant number 1R01AI067913-01A1

\textbf{REFERENCES}


30. Fouset, EM.; Leone, PA.; Ashby, RM. The impact of prostitution and drugs on a rural syphilis outbreak. National STD Conference; San Diego, California. 2002;


36. CDC. Recommendations for Partner Services Programs for HIV Infection, Syphilis, Gonorrhea, and Chlamydial Infection MMWR Early Release. 2008; Vol. 57


Figure 1.
North Carolina counties examined for syphilis outbreak
Figure 2.
A. Incidence of primary, secondary, and early latent syphilis and
B. Number of incident diagnoses of primary secondary, and early latent syphilis
Figure 3.
Columbus County socio-sexual network
Figure 4.
A. Robeson County socio-sexual network and
B. Moore and Montgomery Counties socio-sexual network
Figure 5.
Cumberland County socio-sexual network components with at least four nodes
Table 1

Distribution of selected demographic characteristics by syphilis status of persons documented by Disease Intervention Specialists, North Carolina 1998-2002

<table>
<thead>
<tr>
<th></th>
<th>Infected</th>
<th>Confirmed negative</th>
<th>Unknown</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>(row %)</td>
<td>No.</td>
<td>(row %)</td>
</tr>
<tr>
<td>Total</td>
<td>2,055</td>
<td>(39)</td>
<td>2,497</td>
<td>(47)</td>
</tr>
<tr>
<td>Men</td>
<td>1,056</td>
<td>(36)</td>
<td>1,429</td>
<td>(48)</td>
</tr>
<tr>
<td>Women</td>
<td>999</td>
<td>(43)</td>
<td>1,066</td>
<td>(45)</td>
</tr>
<tr>
<td>County of residence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-outbreak</td>
<td>1,311</td>
<td>(39)</td>
<td>1,530</td>
<td>(45)</td>
</tr>
<tr>
<td>Outbreak</td>
<td>744</td>
<td>(38)</td>
<td>967</td>
<td>(50)</td>
</tr>
<tr>
<td>Columbus</td>
<td>99</td>
<td>(32)</td>
<td>183</td>
<td>(58)</td>
</tr>
<tr>
<td>Robeson</td>
<td>485</td>
<td>(40)</td>
<td>607</td>
<td>(49)</td>
</tr>
<tr>
<td>Montgomery</td>
<td>76</td>
<td>(36)</td>
<td>103</td>
<td>(49)</td>
</tr>
<tr>
<td>Moore</td>
<td>84</td>
<td>(46)</td>
<td>74</td>
<td>(40)</td>
</tr>
<tr>
<td>Race/ethnicity:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>343</td>
<td>(40)</td>
<td>372</td>
<td>(44)</td>
</tr>
<tr>
<td>Black</td>
<td>1,249</td>
<td>(38)</td>
<td>1,585</td>
<td>(49)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>105</td>
<td>(32)</td>
<td>160</td>
<td>(49)</td>
</tr>
<tr>
<td>Native American</td>
<td>252</td>
<td>(39)</td>
<td>317</td>
<td>(49)</td>
</tr>
<tr>
<td>Asian Pacific Islander/Other</td>
<td>13</td>
<td>(46)</td>
<td>12</td>
<td>(43)</td>
</tr>
<tr>
<td>Age on 1 Jan 2000 (year)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-19</td>
<td>139</td>
<td>(31)</td>
<td>292</td>
<td>(65)</td>
</tr>
<tr>
<td>20-24</td>
<td>312</td>
<td>(44)</td>
<td>355</td>
<td>(50)</td>
</tr>
<tr>
<td>25-29</td>
<td>312</td>
<td>(45)</td>
<td>334</td>
<td>(48)</td>
</tr>
<tr>
<td>30-34</td>
<td>345</td>
<td>(49)</td>
<td>313</td>
<td>(44)</td>
</tr>
<tr>
<td>35-39</td>
<td>317</td>
<td>(48)</td>
<td>304</td>
<td>(46)</td>
</tr>
<tr>
<td>40-44</td>
<td>234</td>
<td>(42)</td>
<td>291</td>
<td>(53)</td>
</tr>
<tr>
<td>45-49</td>
<td>158</td>
<td>(46)</td>
<td>168</td>
<td>(48)</td>
</tr>
<tr>
<td>50-59</td>
<td>107</td>
<td>(37)</td>
<td>167</td>
<td>(58)</td>
</tr>
<tr>
<td>≥60</td>
<td>56</td>
<td>(35)</td>
<td>101</td>
<td>(63)</td>
</tr>
</tbody>
</table>
Table 2
Frequency of network component size by distribution network members in components during syphilis outbreaks, North Carolina 1998-2002

<table>
<thead>
<tr>
<th>Component Size</th>
<th>Components (n=907)</th>
<th>Network Members (n=5,299)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>(%)</td>
<td>(Cumulative %)</td>
</tr>
<tr>
<td>2</td>
<td>387 (43)</td>
<td>(43)</td>
</tr>
<tr>
<td>3</td>
<td>217 (24)</td>
<td>(67)</td>
</tr>
<tr>
<td>4</td>
<td>111 (12)</td>
<td>(79)</td>
</tr>
<tr>
<td>5</td>
<td>58 (6)</td>
<td>(85)</td>
</tr>
<tr>
<td>6-10</td>
<td>91 (10)</td>
<td>(95)</td>
</tr>
<tr>
<td>11-20</td>
<td>26 (3)</td>
<td>(98)</td>
</tr>
<tr>
<td>21-32</td>
<td>9 (1)</td>
<td>(99)</td>
</tr>
<tr>
<td>67-100</td>
<td>3 (0.3)</td>
<td>(99+)</td>
</tr>
<tr>
<td>102 &amp; 125</td>
<td>2 (0.2)</td>
<td>(99+)</td>
</tr>
<tr>
<td>232a</td>
<td>1 (0.1)</td>
<td>(99+)</td>
</tr>
<tr>
<td>278</td>
<td>1 (0.1)</td>
<td>(99+)</td>
</tr>
<tr>
<td>895</td>
<td>1 (0.1)</td>
<td>(100)</td>
</tr>
</tbody>
</table>

aThis component refers to a small localized outbreak in Chatham County. It is not considered to be part of the larger outbreak in this analysis, because it occurred in central NC instead of southeastern NC, it peaked and ended during the nascent stage (1998-1999) of the larger outbreak, and none of the Chatham cases had sex partners in the outbreak counties.