Sexual network analysis of a gonorrhoea outbreak

P De, A E Singh, T Wong, W Yacoub and A M Jolly

doi:10.1136/sti.2003.007187

Updated information and services can be found at:
http://sti.bmj.com/cgi/content/full/80/4/280

These include:

References
1 online articles that cite this article can be accessed at:
http://sti.bmj.com/cgi/content/full/80/4/280#otherarticles

Rapid responses
You can respond to this article at:
http://sti.bmj.com/cgi/eletter-submit/80/4/280

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Topic collections
Articles on similar topics can be found in the following collections

- HIV Infection/AIDS (1318 articles)
- Sexually Transmitted Infections (1305 articles)

Notes

To order reprints of this article go to:
http://www.bmjjournals.com/cgi/reprintform

To subscribe to Sexually Transmitted Infections go to:
http://www.bmjjournals.com/subscriptions/
ORIGINAL ARTICLE

Sexual network analysis of a gonorrhoea outbreak

P De, A E Singh, T Wong, W Yacoub, A M Jolly


Objectives: Sexual partnerships can be viewed as networks in order to study disease transmission. We examined the transmission of Neisseria gonorrhoeae in a localised outbreak in Alberta, Canada, using measures of network centrality to determine the association between risk of infection of network members and their position within the sexual network. We also compared risk in smaller disconnected components with a large network centred on a social venue.

Methods: During the investigation of the outbreak, epidemiological data were collected on gonorrhoea cases and their sexual contacts from STI surveillance records. In addition to traditional contact tracing information, subjects were interviewed about social venues they attended in the past year where casual sexual partnering may have occurred. Sexual networks were constructed by linking together named partners. Univariate comparisons of individual network member characteristics and algebraic measures of network centrality were completed.

Results: The sexual networks consisted of 182 individuals, of whom 107 were index cases with laboratory confirmed gonorrhoea and 75 partners of index cases. People who had significantly higher information centrality within each of their local networks were found to have patronised a popular motel bar in the main town in the region (p = 0.05). When the social interaction through the bar was considered, a large network of 89 individuals was constructed that joined all eight of the largest local networks. Moreover, several networks from different communities were linked by individuals who served as bridge populations as a result of their sexual partnering.

Conclusion: Asking clients about particular social venues emphasised the importance of location in disease transmission. Network measures of centrality, particularly information centrality, allowed the identification of key individuals through whom infection could be channelled into local networks. Such individuals would be ideal targets for increased interventions.

A sexual network portrays the sexual inter-relationships within a defined group of people. Sexual network analysis has been used to study the spread of sexually transmitted infections (STIs) within the context of social interactions. It supplies useful theoretical frameworks, methodologies and analytical techniques for traditional contact tracing by identifying specific network members for intervention. This is in contrast with the epidemic curve and analysis of individual characteristics, which are typically used in outbreak investigations but which fail to capture the complexity of connections between cases. Network methods can take advantage of the links afforded by routinely collected STI notification data to determine the spatial and temporal relationships between sexual partners, in both outbreak and non-epidemic settings. Furthermore, analysis of such sociometric risk networks can be complemented by molecular techniques to identify actual and potential routes of infection transmission.

The importance of social context in STI transmission, particularly in marginalised and reclusive populations, was initially recognised by grouping sexually connected people into networks. The applicability of this technique has been demonstrated in studying epidemic transmission of disease. In an outbreak of syphilis, network measures were employed to examine the evolution of sexual links over time and their impact on disease transmission. A combination of ethnographic and social network methods were also used to examine the relative position of people in social, sexual, and drug using networks in determining causes of HIV persistence in the networks.

In this study, we employed sexual network analysis to examine the transmission of Neisseria gonorrhoeae infection in an outbreak. In early 1999, provincial public health staff identified a cluster of gonorrhoea infection in several neighbouring aboriginal (that is, North American Indian and Métis) reservations and industrial towns in the province of Alberta, Canada. In an earlier case-control study of this outbreak we discovered a positive association between gonococcal infection and attendance at a motel bar in the principal town in the region. We were interested in determining whether network methods would corroborate these findings. Here, we describe the nature of sexual networks in the outbreak to determine how risk of infection accords with a network member’s position within the sexual network. We also consider risk in the context of structural differences in a network by comparing a large combined network centred on a social venue with smaller disconnected components.

METHODS

STI services in Alberta include free testing and treatment, provided by primary healthcare providers in medical clinics and hospitals, and by public health nurses in health centres in aboriginal communities. The assay for gonorrhoea detection in the outbreak region was a nucleic acid detection test (PACE-2 GC, GenProbe, San Diego, CA, USA). All suspected cases of gonorrhoea were also tested for genital chlamydia (PACE-2 CT, GenProbe, San Diego, CA, USA).

On receipt of a positive laboratory result for either gonorrhoea or chlamydia, cases and their sex partners are required to be reported to the provincial health department by physicians and laboratories. Public health nurses contact the attending physician for information on reported cases, which they use to locate infected clients, counsel them about their infectiousness, identify potentially infected partners...
within the past 3 months, and refer cases and their partners for treatment.

A computerised surveillance case registry is maintained by the provincial health department, within which each notification of an STI episode for a case is accompanied by information on patient demographics, STI symptoms, treatment received, results of laboratory tests, details of sexual partners identified by the index case, and outcome of partner notification.

**Study population**

Data collection procedures for this study have been previously described in detail. Briefly, people who tested positive for unpenetrated direct gonorrhoea in the outbreak region and were reported to the provincial STI case registry system between 1 January 1999 and 31 December 2001 were included in the study. Three communities (referred to here as communities A, B, and C) experienced 86% of all reported cases. STI notification records of these index cases, along with information about their named contacts (treated and untreated, tested and untested), were extracted from the case registry. Complete demographic and clinical information was unavailable for all sexual partners as public health procedures do not require sexual contacts without laboratory proved infection to be interviewed. In addition to demographic characteristics and STI history, index cases underwent an interviewer administered questionnaire as part of the original case-control investigation, in which they were asked about attendance at public venues suspected of being locations for establishing casual sexual partnerships.

**Analyses**

Univariate comparisons of index cases and their partners were completed in Epi-Info v6.04c (CDC, Atlanta, GA, USA). Sexual networks were constructed by linking subjects with their named partners using PAJEK" and connected components of the resulting networks were extracted. A component is a subgroup of individuals directly or indirectly linked within a network. Components containing four or more such individuals were exported to UCINET V in order to calculate within a network. Components containing four or more such individuals were a subgroup of individuals directly or indirectly linked

networks were extracted. A component is a subgroup of individuals directly or indirectly linked within a network. Components containing four or more such individuals were exported to UCINET V in order to calculate within a network. Components containing four or more such individuals were a subgroup of individuals directly or indirectly linked

Network density represents the proportion of links which exist between all members in a network, out of all those which could exist if each person in the network was connected directly to every other.

The relative importance of each subject in the network was evaluated using three centrality measures: (1) degree centrality, the number of individuals directly connected to a given individual in the network. In partner notification terms, this represents the total number of sex partners named by a client, including those who named the client; (2) betweenness centrality indicates the extent to which individuals act as conduits in the transmission of infection to others within their network. Betweenness centrality is the average number of geodesics or shortest possible paths between a pair of individuals; (3) finally, information centrality estimates the mean distance of a given person from every other person in the network. This measure takes into account the fact that paths other than geodesics can be equally useful for transmission. The combined number of paths between an individual and all other network members are summed using weights proportional to the length of the paths. For example, a direct path between individuals A and B is given a weight of 1, while an indirect path from individual A to C passing through B is given a weight inverse to its length (that is, \(1/(1+1) = 0.5\)). The sum of these weights is averaged for paths between all network members and a given individual in the network is assigned the average information centrality from all weighted paths on which he/she lies.

A k core is a microstructure of a network in which each member is connected to at least two or more other members. It is a measure of the cohesiveness of a subgroup of individuals among whom there are stronger, more direct, or more frequent links (and thus potential transmission routes) than between other subgroups within the same network. In this analysis, we identified two cores in which each member of a subgroup was connected to at least two other members in the same subgroup. By calculating the above three centrality measures and two cores, we intended to identify individuals who were most central in their network and likely to be most influential in disease transmission. Firstly, network measures of subjects directly linked to the bar, which was the suspected source of infection spread, were compared to those of members on the periphery of the networks. Secondly, we compared such subjects with network members who had lower centrality measures in order to define sociodemographic and behavioural risks among those holding key positions in their network. Thirdly, we compared isolate cases—namely, individuals who identify no sex partners and who themselves are unnamed by others—with non-isolate cases. We postulated that the former may represent higher risk individuals who intentionally remain anonymous in casual partnering, while the latter are unwilling or unable to name their sexual partners and who presented early on in the outbreak. The linkage of sexual partners showed wide geographic distribution; sexual contact was reported between people residing as far as 220 km apart. Sixty nine network components were found, ranging in size from one to 39 individuals. Measures of size and component structure are shown in table 2. Components with radial structures were mainly centred on one or two individuals, with links radiating from them in a star shape, whereas linear networks were predominantly links between one person and the next in a long line. The density of the networks decreased with increasing network size, indicating that individuals in the largest networks had the lowest proportion of all potential links to others in their network.

As expected, mean degree centrality and betweenness centrality increased with larger network size. The largest networks contained individuals with a higher average number of partners and who were more likely to lie on a larger number of geodesics than those in smaller networks. In contrast with degree and betweenness centrality, mean information centrality decreased with increasing network size. Given that information centrality is a measure of the quality of the information decreases and the amount of “noise” increases as the number of paths and their lengths increases from the source to recipient. The two core

**RESULTS**

In all, 182 network members, composed of 107 index cases of gonorrhoea (not counting seven re-infections) and 75 named sexual contacts were identified in the outbreak networks. Table 1 shows the sociodemographic profile of study subjects. Index cases ranged in age from 15 to 60 years and 58 (54%) were women. Of the 75 sexual partners, 43 (59%) were men and 32 (43%) had confirmed gonorrhoea infection.

Response rates of 35% for men and 62% for women were reached in index subjects who were asked about patronage at public venues. Of the 17 male and 36 female index cases who responded to this query, 35% of men and 33% of women reported a motel bar in the principal town in the region (community A) as the place most frequently visited.

The linkage of sexual partners showed wide geographic distribution; sexual contact was reported between people residing as far as 220 km apart. Sixty nine network components were found, ranging in size from one to 39 individuals. Measures of size and component structure are shown in table 2. Components with radial structures were mainly centred on one or two individuals, with links radiating from them in a star shape, whereas linear networks were predominantly links between one person and the next in a long line. The density of the networks decreased with increasing network size, indicating that individuals in the largest networks had the lowest proportion of all potential links to others in their network.

As expected, mean degree centrality and betweenness centrality increased with larger network size. The largest networks contained individuals with a higher average number of partners and who were more likely to lie on a larger number of geodesics than those in smaller networks. In contrast with degree and betweenness centrality, mean information centrality decreased with increasing network size. Given that information centrality is a measure of the quality of the information decreases and the amount of “noise” increases as the number of paths and their lengths increases from the source to recipient. The two core
measure revealed nine individuals in the largest component (n = 39) who were directly connected to at least two other members within the subgroup, forming a circle of potential transmission (fig 1). Comparisons of degree centrality and age were not significant, nor was there any observed transmission (fig 1). Comparisons of degree centrality and betweenness centrality revealed nine individuals in the largest component of the two core and other members of the same component, or between members of the larger networks containing three or more individuals (not shown).

As shown in table 3, individuals who reported patronising the motel bar had significantly higher information centrality (p = 0.05) and were older (p = 0.04) than non-patrons within each of their components. By virtue of common patronage at the bar, a sexual network was constructed that joined all bar patrons belonging to different components using the bar as the central node (fig 2). The resulting large component contained 89 people representing 49% of the study population, including 32 subjects who were positive for gonorrhoea; counts for subject characteristics may not sum to total because of missing data.

DISCUSSION
In 1985, Potterat et al demonstrated the importance of social venues where sex partners meet. Since then, the inclusion of social venues has been inconsistent in STI investigations. However, the current network study and the previous case-control study of this outbreak re-emphasise the notion that...
social venues, such as the motel bar in this study, may be essential and, in this case, the only factor defining individuals at risk of infection. Secondly, our findings confirm that sexual network analysis can allow the linkage of seemingly isolated outbreaks from geographically distant communities. This notion also emphasizes the fact that targeting individuals in isolation, rather than considering their complete sexual space and geography, may limit the impact of interventions. On a practical note, we confirm that asking clients about local "pickup joints" is successful in delineating the extent and pattern of gonorrhoea transmission within sexual networks, while allowing access to a greater proportion of the population at risk. This may be because some STI clients are more willing to name places rather than people with whom they have had sexual contact. The naming of partnering locations enhances patient recall and can concurrently improve contact tracing yield.

To our knowledge, this is the first empirical study to demonstrate conclusively that people at highest risk of infection (that is, those who attended the bar) also hold

**Table 3** Characteristics and centrality measures of gonorrhoea index cases who were bar patrons compared with non-bar patrons, belonging to components of size 4 or greater

<table>
<thead>
<tr>
<th></th>
<th>Attended bar (n = 9)</th>
<th>Did no attend bar (n = 18)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2 (22.2)</td>
<td>5 (27.8)</td>
<td>1.0</td>
</tr>
<tr>
<td>Female</td>
<td>7 (77.8)</td>
<td>13 (83.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (interquartile range)</td>
<td>24 (22–32)</td>
<td>20 (16–26)</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Residence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community A</td>
<td>1 (11.1)</td>
<td>0 (0)</td>
<td>0.15</td>
</tr>
<tr>
<td>Community B</td>
<td>5 (55.6)</td>
<td>14 (77.8)</td>
<td></td>
</tr>
<tr>
<td>Community C</td>
<td>3 (33.3)</td>
<td>2 (11.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1 (11.1)</td>
<td>0 (0)</td>
<td>0.33</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>8 (88.9)</td>
<td>18 (100)</td>
<td></td>
</tr>
<tr>
<td><strong>Infection characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-infection with chlamydia</td>
<td>4 (44.4)</td>
<td>8 (44.4)</td>
<td>1.00</td>
</tr>
<tr>
<td>Re-infection with gonorrhoea</td>
<td>1 (11.1)</td>
<td>1 (5.6)</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Centrality measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median degree centrality (range)</td>
<td>2 (1–3)</td>
<td>1.5 (1–2)</td>
<td>0.43</td>
</tr>
<tr>
<td>Median betweenness centrality (range)</td>
<td>3 (0–19)</td>
<td>1.5 (0–37)</td>
<td>0.19</td>
</tr>
<tr>
<td>Median information centrality (range)</td>
<td>50 (36–63)</td>
<td>35 (29–50)</td>
<td>0.05</td>
</tr>
</tbody>
</table>
the highest information centrality measure in their own networks. In a study of concurrent sexual, drug sharing, and social networks in Colorado Springs, eight measures of centrality (including the three used here) were used to arrive at similar epidemiological findings as those found in the current study, although without the benefit of a case-control study. The advantage of using information centrality is that it accounts for the probability of transmission from (1) directly adjacent members who are partners of the individual (that is, degree centrality), (2) all members connected to him/her via geodesics (that is, betweenness centrality), and (3) all other individuals in the network regardless of how small their contribution might be. In this sense, information centrality most closely resembles the complexity of real relations when compared with the simpler measures of betweenness and degree centrality. Furthermore, subjects with high centrality connected larger populations with their local networks as a result of their links to the bar, and were thus also influential in transmitting disease. No clinical or demographic differences were found between bar patrons and non-patrons. However, psychological or sociological factors such as extroversion and social role may predispose the most central network members to positions of influence, both in terms of transmission of infection—but more importantly—in the dissemination of preventive information for which they should be placed at highest priority.

In addition to the bridges, the two cores provide a circular structure which allows for continuous cycling of infection, despite the occasional removal of people during treatment. Also, the two core allows for each individual to be infected by either of the people connected to him/her, indicating a level of pre-emption conforming to the earliest definition of “core group.”

Given that partner notification information tends to be incomplete, this results in unknown or inaccurate transmission pathways. While the importance of such missing links to the overall network is difficult to assess, forgetting to name sex partners has been found to be arbitrary; the only caution being that individuals with the most sex partners tend to also forget the most. However, the existence of a sexual partnership, once reported, makes misclassification of network members belonging to certain subgroups unlikely. A higher incidence of unlinked, isolate cases was noted later in the outbreak rather than earlier, contrary to our expectation. As the outbreak progressed and was better publicised, it is possible that individuals who suspected themselves to be infected became more inclined to present independently for care. Also, isolate cases were unevenly distributed in the communities, possibly because of different methods and personnel involved in partner notification. The direction of transmission in this outbreak could not be ascertained as a result of insufficient data on time of infection and lack of information on gonococcal strains to confirm linkages. The availability of genotyping data, which is not permitted by the

## Key messages

- This is the first empirical study to provide evidence of an association between risky sexual behaviour and relative position in a sexual network, as measured by information centrality. People with the highest information centrality scores were the most influential within their network in transmitting gonococcal infection. This differs from the traditional view that those who have risky behaviours, such as a large number of partners, are more likely to transmit disease. From a public health perspective, people with a central role in their network as defined by information centrality should be targets of intervention and can serve as facilitators in the dissemination of health information.

- This study shows that the traditional case-control investigation of an outbreak and a network analysis can arrive at similar results (that is, that the “source” of infection is the bar), the implication of which is that the cheaper and easier network analysis may be substituted for the more complex case-control analysis.

- As demonstrated by previous studies, sexual networks can allow linkage of seemingly isolated outbreaks and identification of potential core group members from geographically distinct communities.

- The study confirms that social venues should continue to be recognised as important interchanges of disease transmission and be queried about when collecting exposure information from STI clients.

### Table 4
Comparison of gonorrhoea index cases belonging to networks (connected) and those not reporting sexual partners (isolate)

<table>
<thead>
<tr>
<th></th>
<th>Connected (n = 156)</th>
<th>Isolate (n = 26)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>80 (51.3)</td>
<td>13 (50.0)</td>
<td>0.90</td>
</tr>
<tr>
<td>Female</td>
<td>76 (48.7)</td>
<td>13 (50.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (interquartile range)</td>
<td>30 (17–33)</td>
<td>24 (19–28)</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>8</td>
<td>3</td>
<td>0.55</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>124 (79.5)</td>
<td>23 (88.5)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Residence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community A</td>
<td>16 (10.3)</td>
<td>7 (26.9)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Community B</td>
<td>54 (34.6)</td>
<td>3 (11.5)</td>
<td></td>
</tr>
<tr>
<td>Community C</td>
<td>47 (30.1)</td>
<td>8 (30.8)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>19 (12.2)</td>
<td>8 (30.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Time to diagnosis from start of outbreak (days)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (interquartile range)</td>
<td>47 (23–69)</td>
<td>93 (49–101)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Infection characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-infection with chlamydia</td>
<td>34 (21.8)</td>
<td>7 (26.9)</td>
<td>0.57</td>
</tr>
<tr>
<td>Re-infection with gonorrhoea</td>
<td>7 (4.5)</td>
<td>0 (0)</td>
<td>0.60</td>
</tr>
</tbody>
</table>
current molecular diagnostic technique, would have strengthened the purported links between subjects. In addition, information on timing of partnerships and precise onset of symptoms would allow for a more chronologically accurate analysis of transmission.

In conclusion, more information about infection risk was made available by asking cases about particular social venues they patronised, emphasising the relevance of locations in disease spread. The hypothesis that apparently isolated outbreaks were connected was supported by viewing geographically distant sexual partners as belonging to common sexual networks. We demonstrated that individuals in key positions within their respective networks, measured by information centrality, may be important in channelling infection into local networks. Finally, as both the traditional case-control investigation and the current network analysis of this outbreak arrived at the same conclusion (that is, that the bar was implicated as the “source” of infection), network analysis may in the future represent a cheaper, easier and yet quantitative alternative to the case-control study, as the former does not require additional interviews of control subjects.

ACKNOWLEDGEMENTS
The authors thank the regional STI nurses who participated in the coordination of data collection.

This work was presented in part at the 15th meeting of the International Society for Sexually Transmitted Disease Research and at the 13th Sunbelt Conference on Social Network Research.

CONTRIBUTORS
PD and AJ contributed to the study design, data collection and analysis, and preparation of the draft manuscript; AS, TW, and WY were responsible for conceptualisation of the study; all authors took part in revising the manuscript.

Authors’ affiliations
P De, T Wong, A M Jolly, Centre for Infectious Disease Prevention and Control, Health Canada, Canada
A E Singh, Office of the Provincial Health Officer, Alberta Health and Wellness, Canada
W Yacoub, First Nations and Inuit Health Branch, Health Canada, Canada

T Wong, A M Jolly, Department of Epidemiology and Community Health, University of Ottawa, Canada
Conflict of interest: None declared

REFERENCES
12. Batagelj V, Mrvar A. PAJEK—program for large network analysis (vlado.fmf.uni-lj.si/pub/networks/pajek/).