The relevance of social network analysis on the epidemiology and prevention of sexually transmitted diseases

Importância da análise de redes sociais na epidemiologia e na prevenção das doenças sexualmente transmissíveis

Abstract

Historically, the epidemiology of sexually transmitted diseases (STD) has been based on individual attributes and behavior. However, STD constitute a good example of diseases that depend on personal contacts for dissemination. Social network analysis is a relatively new technique that studies the interactions among people. Since 1985 when it was first used for STD, some studies have been done using the technique, especially in the last 10 years. The two network-based designs, sociocentric or complete networks and egocentric or personal networks, are currently recognized as important tools for a better understanding of STD's dynamic. Here an overview is presented of social network analysis: the technique, its use, and its limitations. Ethical considerations regarding social network analyses are also briefly discussed.

Sexually Transmitted Diseases; Communicable Disease Prevention; Community Networks

Introduction

According to the World Health Organization (WHO), there were approximately 340 million new cases of sexually transmitted diseases (STD) worldwide in 1999. If just the prevalence of curable STD, such as gonorrhea and chlamydia, are included, there were an estimated 116.5 million cases. There are no data available for Brazil, as these diseases do not require compulsory notification. WHO data for Brazil indicate that there are around 12 million new cases of STD per year. A study carried out at the Serviço de Dermatologia at the Santa Casa da Misericórdia do Rio de Janeiro [Dermatology Department, Rio de Janeiro Mercy Hospital] between January 1999 and January 2002 found that among the 395 patients examined, there was a prevalence of 13% for syphilis and 43% for condyloma acuminata, among other STDs evaluated (data not shown). Taking into consideration that these figures may be underestimated, it can be concluded that, in spite of the efforts being made to control and prevent STDs, they are, in general, spreading worldwide, bringing with them all the associated health problems. In this context new epidemiological techniques that aid in the prevention of STDs could be a great help in the attempt to control these diseases.

Since 1985, when Klovdahl presented his study which is considered to be the first one to employ the analysis of social networks in the
study of STD, various other researchers have
dedicated themselves to this technique with the
object of encountering new tools to help in the
understanding of the dynamics and prevention
of STD. There have been various studies analyz-
ing the social network and demonstrating its rel-
ance for STD research, including those already
carried out, and ongoing studies, especially in
the United States. Recently, in Brazil, Barbosa et
al. 2 carried out a review of the subject, probably
the first group to approach it in Brazil. This article
will attempt an overview of the subject, giving
a technical description, and describing its uses
and current applications as well as its limitations.
Ethical considerations related to social network
analysis will also be touched on briefly.

**Analysis of the social network**

Historically STD epidemiology has been based
on individual attributes and behaviors, attempt-
ting to identify those who could get an infection.
However, a basic premise of STD is that a person
without a sexual partner does not acquire it. In
other words contact between people is the most
basic factor for transmission to occur. Epidemi-
ological techniques normally employed for study
of STDs do not take into account in their analy-
thesis the contributions that interactions between
people have on the probability of dissemination
of STDs; only individual risks for acquiring them
are analyzed.

A social network consists of a series of indi-
viduals or groups connected by links that repre-
sent some kind of relationship (e.g. friendship)
or interaction (e.g. sexual) 3. The investigation
of social networks permits an evaluation of the
influence that the connections between people
have in the transmission of a given disease. In
this type of study, information about each indi-
vidual, or, in the terminology of the social net-
work, about characteristics related to egos, is col-
lected. Afterwards, information is collected about
those people that have some kind of relationship
with the egos. These individuals constitute the
neighborhood of each ego. The concept of neigh-
borhood employed in social network analysis is
graph theory, where adjacent points are points
connected by a line and all the points that one
point is connected with constitute a neighbor-
hood. In this definition there are no specific
types of interaction that are going to depend on
a scientific question to be answered. In the case
of sexual networks, the social network analysis
could include regular partners, casual partners
or anonymous ones; all are representatives of
the ego's neighborhood. Social network analy-
sis can also include other types of relationships
for analysis, like friends and acquaintances, and
try to identify the role these relationships play
in the dynamics of infection by STD/HIV. The
information collected in respect to partners is
based, in general, solely on the perceptions of
the ego in respect to its partners. Alternatively,
the partners could be contacted, so the veracity
of the information collected about them could
be confirmed. Intuitively it can be inferred that
the risk of a person acquiring an STD is related
to contacts with their neighborhood, which re-
forces the importance of this type of study in
the prevention of STDs. The next logical step in
the analysis is to start studying the connections
between diverse personal networks, so forming
a sociogram. On analyzing a much larger field
of interactions between people a better idea can be
obtained of the direct and indirect risks for some-
one acquiring a STD. Aside from this, the global
structure of the sexual network (sociometric net-
work) is fundamental to a better understanding
of the dynamics of transmission of an infectious
agent that depends on intimate contact (sex) for
contagion.

Based on this explanation, two types of social
network analysis are defined: the egocentric or
personal, and the sociometric network. With the
first type it is possible to study the relationship
between the ego and its neighborhood, all infor-
mation collected, generally, centered on the ego.
With the second type of analysis it is possible to
study the totality of individuals in a determined
population, and all the links between them. Each
of the two types described has its own respec-
tive utility and complexity. Study data from the
eyecentric network are more easily collected as
they can be gotten when doing other research.
For composing sociometric networks, the data
ideally should be collected as part of research
specifically designed to this end, as it possesses
still other complexities, such as the definition
of population to be studied, ethical problems in
the search for partners, and problems with data
analysis. In regard to data analysis, in the case of
eyecentric networks, it is basically descriptive,
as much in the data that refer to the ego as those
that refer to the neighborhood. Complementing
this, measurements identified as important in
the dissemination of a STD in a determined pop-
ulation, such as the degree of concomitant stable
relationship, and the level of mixing between
people in a population, can also be evaluated
utilizing data from egocentric networks 4. For the
studies of sociometric networks where there is an
immense degree of interdependence among the
diverse links existant, traditional statistics cannot
be utilized. Alternatively analysts use graph the-
Dynamics of STD

It is natural that the concept of social network be incorporated into the study of STD dynamics. The probability of acquiring or transmitting STDs not only depends on an individual's serology and risk behaviors, but also on the characteristics of the social network around the person. In other words, how a person interacts in a group or given population. In this sense, couples that have a strictly monogamous relationship do not become infected, even though they never use condoms during sexual relations. On the other hand, if, for example, in a conjugal relationship the husband has various sexual partners, his wife will have a great risk of becoming infected. Meanwhile, the wife will be protected again from her “high-risk husband” if he uses condoms in all his extra-marital sexual relations. The individual's risk of contracting an infection depends on the behavior of the partner of that person, the position of their partners on the sexual network and STD prevalence on the network.

The concept of core individuals or groups has been much debated in the literature since its description. These can be defined as groups or individuals with greater probability of infecting other people, although there is no standard definition or consensus in the studies. What importance do individuals or core groups have in transmission of STD? Rosenberg et al. demonstrated that the proportionately high syphilis transmission rate in populations of some cities in Louisiana, in comparison to other areas of the United States, could have been maintained by a relatively small number of individuals with a high rate of STD and high number of new partners (their definition of core individuals). If they had had a central position in a sexual network. In the case of groups of core people, two other parameters are important: the concept of links between populations, or bridging groups; and a standard of mixing in a given population, or mixing pattern. A bridge can be defined as a person or group of people that function as transmitters of an infection from a population with high prevalence to another with less prevalence. Lowndes et al. showed the importance of these two concepts, central group and bridge, in the dynamic of STD/HIV transmission. They described females as central groups responsible for STD transmission. The authors showed that women in the general population with low risk behavior in general had a high risk of being infected by their male partners who, as mentioned above, served as the bridge between the high risk group of sex workers and their wives. Another study carried out by Morris et al. in Thailand, demonstrated the importance of bridges in transmission. Now the standard of miscegenation can be described as the degree of interaction between people within a risk group or the degree of interaction between groups under risk. The importance of core groups in the transmission of an STD depends on the degree of interaction between those groups and the peripheral, or non-core, groups. If, for example, the people in the core groups (e.g. injection drug users) tend to interact, or mix, only within their own group, then the infection will have a tendency to stay confined in their group only. If, meanwhile, the core groups have a tendency for mixing with the non-core groups then the infection will have a greater potential to be distributed in a population and a greater prevalence in it. Kottiri et al. utilized social network analysis to attempt to explain why there was a greater prevalence of HIV among black injection drug users compared with white injection drug users. The authors' objective was to evaluate the social network to see if it could explain this variation. The study was carried out in Bushwick, New York and the authors concluded that the sexual networks and injection drug networks were homogeneous in relation to race/ethnicity. The authors came to the conclusion that as the probability of a black person to be infected with an STD was greater, due to a greater prevalence of STDs among African-Americans, and as they tended to have interactions among themselves, this standard of miscegenation could be responsible for maintaining a higher level of rates among people of African-American descent. Some authors prefer using the term “core networks” instead of individuals or core groups. These authors argue that what is important in the transmission of an STD is the structure of the sexual network of the population. Potterat et al. analyzed cases of chlamydia in Colorado Springs, Colorado, United States during a four year period, to verify whether the structure of the sexual network in this population could indicate the phase of the epidemic in the city.
The authors showed that the sexual network was fragmented and dendritic, and that, because of this, it had a low infection transmission potential. On the other hand, sexual networks of the dense type, with various connections between their components (highly connected) are responsible for a high potential of infection transmission, as was demonstrated by the authors in their comparison with sexual network structure of an outbreak of STD in the same city between 1990 and 1991, among street gang members using cocaine. This fact had already been analyzed in Atlanta by Rothenberg et al.\textsuperscript{14} in the transmission of syphilis. Another article confirmed the influence that the sexual network structure can have on the transmission possibility in a set population. Woodhouse et al.\textsuperscript{15}, studying heterosexual individuals at high risk of acquiring HIV, again in Colorado Springs, showed that the majority of people that became infected with HIV where located in small subgroups (or components, in the terminology of social network analysis), which suggested that HIV-positive individuals were socially isolated from the other people in the network. This fact suggests that, although this was a population at high risk for acquiring HIV, the spread of the infection did not occur because the network structure of the group served as a barrier against transmission.

Much of what has been published related to social network analysis refers to research with sociometric networks. Some studies have also been developed in egocentric networks. Meanwhile it is always good to emphasize the importance of studying personal networks due to the relative ease of carrying out these studies, making possible responses to scientific questions about the their epidemiological impact on STD/HIV. An example of this is an article published by Ghani et al.\textsuperscript{16}, in which the authors demonstrate that the risk of a person acquiring an infection is determined by the structure of his/her personal network.

The study of STD epidemics

Another use for social network analysis was introduced at the same time this technique was used for the first in the first study of STD. Klovdahl\textsuperscript{17} used data related to the first reported cases of HIV collected by the CDC (Centers for Disease Control and Prevention) to demonstrate the utility of social network analysis in the evaluation of epidemic outbreaks. After putting together the sociogram of the first cases, the author was able to demonstrate various aspects of the emerging epidemic. In this case social network analysis was able to give support to the theory that an infectious agent was involved in the etiology of AIDS, principally because the contacts between the people of the network suggested transmission through an agent (sexual, in this case). In the article the author also discussed themes related to the influence of the network structure on the disease transmission of a population development of strategies to stop transmission and long-term changes in social networks (interactions are dynamic and change over time), themes that were only much later acknowledged and investigated.

In 1999 the CDC and the Secretary of Health of the State of Mississippi applied social network analysis to the study of an outbreak of HIV among adolescents in that state\textsuperscript{18}. In this study both sexual and social contacts were tested and while both sexual and social partners were similar in relation to numerous variables, one variable was different in the two groups: younger women maintained sexual relations with older men in the sexual network. This mixing pattern was found to be one of the causative factors of the outbreak.

In 2004, De et al.\textsuperscript{19} described the use of social network analysis in the investigation of an outbreak of gonorrhea in Alberta, Canada. The authors were able to show that some of the measurements utilized in the social network analysis could be useful in the identification of people that served as carriers in the dissemination of STD to local networks. These people were not necessarily the central individuals previously described, but were people with a high degree of information centrality. This measurement refers to the average distance from a person to any other person in the network, or, in other words, measures the degree of proximity of an individual in relation to all the people in the network. In theory, if this person becomes infected, the dissemination of infection agent would be earlier due to his/her proximity to all other members in the network. The authors further affirm that this fact will imply another focus of prevention, issuing from the central individuals (e.g. with various partners) to individuals with a great proximity to all other people in the network (high degree of information centrality).

Social network analysis and geographic analysis

More recently some authors have begun to recognize the importance of geographic location of social networks which can define, for example, locals or core areas in the transmission of an STD instead of individuals or groups. Wylie & Jolly\textsuperscript{20}
utilized social network analysis and geographic analysis in Manitoba, Canada with the objective of describing the model for transmission of chlamydia and gonorrhea in that city. Various components were described inside the sexual network that varied in size (number of people). The analysis of the largest component, consisting of 82 people, demonstrated that individuals were rarely in group. On the contrary, this component showed a high degree of dispersion of its members around the city. Furthermore many components contained people that were geographically located outside areas with high incidence. Meanwhile, these people possessed multiple partners, which made re-entry of the infection into core areas possible. The evaluation of the components demonstrated the potential for STD transmission between isolated rural communities and between distinct areas in the urban landscape. Wylie et al. published a study carried out in Manitoba, Canada where the authors utilized molecular techniques to identify possible geographic cluster areas for chlamydia infection. The authors also used social network analysis to evaluate if the cluster areas corresponded to the individual sexual networks. There were ten geographic cluster areas identified, based on chlamydia genotype data. The clusters were isolated without overlapping areas. The analysis combined the data (geographic, molecular, and social network) to show that each geographic cluster seemed to represent a network with a distinct transmission model. Rothenberg et al. sought to examine the relationship between social distance and geographic distance. The data were collected in Colorado Springs, between 1988 and 1992, from people with a high risk of acquiring HIV. At the end of the study the authors were able to conclude that there was an association between social distance and geographic distance, and that this association, in a group of people at risk to acquire HIV, demonstrated the importance of geographic clusters in the transmission of the disease.

Use in epidemiology and prevention of STD

One of the initial campaigns for the prevention of AIDS in Brazil was based, probably unintentionally, on the idea of social networks. Having as its base the poem *Quadrilha* by Carlos Drummond de Andrade, the campaign divulged the idea that people in contact with their partners are equally in contact indirectly with their partners’ partners, with their partners’ partners’ partners, and so on. The message was clear: a person that loves another person, that loves a third, that loves someone that has died of AIDS, also is at risk of contracting HIV. In the United States, Parran described, in 1937, a venereal disease program that was used in the country for various decades. One of the pillars of this program was the notification, search, and treatment, of partners of infected people, which in social network analysis are defined as egocentric data. Therefore, it can be seen that even though the concept of social network has been utilized in public health for some time, it is only recently that social network analysis has been recognized as playing an important role in the epidemiology and prevention of STD.

In his article published in 2001, Rothenberg describes various forms of intervention based on social network analysis. The breakup of the network structure through which transmission occurs, which reminds one of the beginning of the AIDS epidemic, when a prohibition of saunas and other places where HIV transmission occurred was suggested. This proposal, which was based on the idea of removing key areas of transmission from the network structure, had quite a negative impact at the time it was proposed, and is not feasible for application today. Core people in a social network can be utilized in two ways: as they are socially influential, they can be used to diffuse prevention information in their social network; on the other hand, because they are the core of a sexual network, they can become targets of preventative measures with the goal of impeding transmission of the infection to their partners. The use of people as recruiters of people at risk on their social network is another form of intervention based on social network analysis proposed by the author. Although the author has proposed other forms of intervention, only the three ideas mentioned above will be cited as there are already studies that demonstrate their potential for STD prevention.

Michaud et al. demonstrate an example of intervention based on locales where sexual partners are encountered. The authors affirm that the method based on actions done in central areas is unique, as it has as its target not the individuals at risk, but the areas where people infected with syphilis say they encountered sexual partners. The interventions are usually long-term in these locales and, in the case of the study cited, based on the offer of a test for syphilis and HIV by a trained professional, in addition to treatment in case of positive cases with notification of partners. Using these methods, the authors were able to identify people positive for syphilis that had not been identified in traditional investigations.

In 2001, Liljeros et al. concluded in the journal *Nature*, that individuals with multiple sexual partners would be perfect targets for safe
sex campaigns to prevent STD. As they were able to demonstrate, the majority of people have few sexual partners, while a few individuals have several sexual partners. In other words, the structure of the sexual network, in this case, is strongly dependent on some key individuals, as was mentioned earlier. This is called a scale-free network. The importance of this finding, according to the authors, is the fact that the structures of the scale-free network are highly susceptible to the removal of the most well connected individuals (people central to the sexual network). Schneeberger et al. examined data from four population studies, three in England, and one in Zimbabwe, and recorded the existence of scale free network in the sexual partners of these populations.

If central individuals are important in the spread of an infection, the same can be said about their importance in disseminating information about the prevention of STD. The theory of diffusion introduced by Everett in 1962, affirmed that new ideas, practices, or innovations enter into a community through an external source. The adoption of an innovation flows, therefore, through interpersonal contacts on a network. Various factors influence the speed of diffusion, among them; the relative advantage to what is being used at the moment, cost, and network structure. Individual relations can explain someone's adoption of an innovation (egocentric level), while the network structure and the position of the individual on it can explain the speed of the diffusion and its final reach. Two studies can be found which utilize leadership training in their communities (populations) to diffuse preventive measures against HIV, both showing results that apparently favor the strategy, one in India and the other one in Russia and Bulgaria.

In June 2005 the CDC disseminated preliminary data from a pilot study started in seven cities in the United State. to evaluate an HIV prevention strategy based on social networks. Using this technique, people testing positive or negative (but still at risk) for HIV are required to recruit their social, sexual, and illicit drug partners (called network associates) that they believe are at risk for acquiring HIV, for HIV testing and counseling. There were 130 people who became recruiters, with a total number of 814 network associates recruited, of which 90% were recruited by the recruiters themselves, and 10% by health professionals, based on data furnished by the recruiters. Some groups like injection drug users were more effective recruiting associates, while others, like male homosexuals were more effective in recruiting people that came and tested positive for HIV.

**Limitations**

While the analysis of social networks is fundamental to the understanding of the dynamics and prevention of STD, these studies should be evaluated very cautiously. As with any design of epidemiological study, many limitations should be mentioned. However only three limitations, of great importance to social network analysis, will be described. Some studies indicate that, when asked to remember all the partners with whom they have had some sort of interaction, the network participants not only tend to forget partners, but also tend to not report some partners with whom they interacted. This fact could create a bias in social network studies, seeing that the contacts that are not reported do not represent a random sample of the population. Some methods have been proposed to improve the recall of partners in social network studies. The localization of the partners reported by a case index can also be a problem in social network studies. The localization of sexual partners in well-established STD programs in the United States is calculated at around 50% of all the partners exposed. Another limitation is more important in egocentric network studies, where the researchers generally have to believe in the information given through the index case about the risk factors of their partners. In an attempt to study the veracity of the information of a case index, Neaigus et al. evaluated the so-called dyads (e.g. husband/wife). These authors were capable of showing that the case index is reliable enough in giving information about their partners such as: age, sex, race, intravenous drug use, sexual relations, and condom use.

**Ethical considerations**

Strategies to contain outbreaks of some diseases are based on the identification and treatment (or use of prophylactic medicine) of intimate contacts from the case index, like, for example, in the case of the meningococcal disease. A basis for this strategy can also be found in social network analysis, which is quite effective when well applied. Two basic differences between STDs and other pathologies are found in the greater incubation periods for STDs, which generates a greater number of contacts (can be extremely high), and ethical problems related to the search for sexual partners, such as domestic violence and confidentiality. These two factors, aside from another factor related to Brazilian health professionals' lack of experience, for various reasons, in
the matter of carrying out contact searches, can be identified as crucial in the debate about the introduction of social network analysis in Brazil.

Perhaps the greatest ethical problem in social network analysis is related to contacting partners indicated by case indexes, and the possibility that these contacts could bring domestic violence problems. The main focus of discussion is related to confidentiality. However, other problems such as consenting of partners to participate in a study, have also been debated. According to North & Rothenberg, the association between domestic violence and notification of partners has been little studied, but presents a real problem. The CDC also recommend that a hostile reaction by a partner be anticipated by the health professional as well as the case index, in the best way possible in all visits to health centers.

Heated discussions have taken place about the ethics of indirect recruitment of people in studies of the sexual network, mainly in regard to the relationship between researchers/research and ethics committees. The main focus of discussion is related to confidentiality. However, other problems such as consenting of partners to participate in a study, have also been debated. According to Margolis, indirect recruitment is a strategy that diminishes respect and privacy and depends on the desire of an individual to violate the confidentiality established in a relationship.

Nevertheless Klovdahl, argues that there are ways for researchers to guarantee a high level of confidentiality in the information collected in social network studies, however these methods can be, in general, very costly. This author argues in his article that even though a study of social network possesses certain particularities in its planning and execution, these studies are not incompatible with secure practices of human beings recommended by ethics committees.

Discussion

Various other studies have been encountered in this search, in medical and sociological literature. This article includes only those that are judged most important for a brief introduction to the vast world of social network analysis. This article searches to define social network analysis and describe its principal uses as an epidemiological technique and in the prevention of STD, as well as its limitations and ethical problems.

To finalize it is important to note that, in writing this article, no study of sexual networks that uses real Brazilian data was found. It is believed that this technique will be of great utility in the study of STD in Brazil, and, because of this, studies of social networks should be carried out here.

Resumo

Historicamente, a epidemiologia das doenças sexualmente transmissíveis (DST) tem se baseado em atributos e comportamentos individuais. Entretanto, tais patologias se configuram em um bom exemplo de doenças que dependem do contato pessoal para disseminação. A análise de rede social é uma técnica epidemiológica que estuda as interações entre as pessoas, baseando-se, no caso das DST, na busca de contatos. Desde 1985, quando foi utilizada pela primeira vez no estudo das DST, várias pesquisas têm sido realizadas utilizando-se dessa metodologia, principalmente nos últimos dez anos. Suas duas formas de uso, análise sociométrica ou completa e análise egocêntrica ou pessoal, são atualmente reconhecidas como peças importantes para um melhor entendimento da dinâmica das DST. Em nosso artigo, procuraremos apresentar uma revisão do assunto no que diz respeito à descrição da técnica, suas utilidades e aplicações corriqueiras e suas limitações. Discutiremos também brevemente, as considerações éticas relacionadas à análise de rede social.

A. R. S. Périsé and J. A. C. Nery participated in all the stages of literature review, design, writing, and editing the text.

Acknowledgements

The authors of this article wish to thank the Fogart International Center for financial support to one of the authors (A.R.S.P.) through a Doctorate training scholarship (AIDS International Training and Research Program/AITRP, National Institutes of Health Research Grant D43-TW001041).
References


Submitted on 15/May/2006
Final version resubmitted on 06/Oct/2006
Approved on 20/Oct/2006