Advances in research of HIV transmission networks

Zhi-Long Dong¹, George Fu Gao¹, Fan Lyu²

¹National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; ²National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China.

Abstract
Transmission network analysis is a crucial evaluation tool aiming to explore the characteristics of the human immunodeficiency virus epidemic, develop evidence-based prevention strategies, and contribute to various areas of human immunodeficiency virus/acquired immunodeficiency syndrome prevention and control. Over recent decades, transmission networks have made tremendous strides in terms of modes, methods, applications, and various other aspects. Transmission network methods, including social, sexual, and molecular transmission networks, have played a pivotal role. Each transmission network research method has its advantages, as well as its limitations. In this study, we established a systematic review of these aforementioned transmission networks with respect to their definitions, applications, limitations, recent progress, and synthetic applications.

Keywords: Human immunodeficiency virus; HIV; Transmission networks; Social network; Sexual network; Molecular transmission network; Molecular epidemiology

Introduction
Network analyses provide an approach for the application of analytical insights and computing techniques to social support, social impacts, and infectious diseases study flows within populations and organizations. Considering the entire population as a group of interconnected individuals forming a large social network provides a more comprehensive understanding of the propagation of certain infectious diseases. In the early days of transmission network research, many studies had already started implementing social networks to explore the characteristics of the human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) epidemic, as well as the corresponding prevention strategies. HIV transmission generally occurs through risky behaviors, including close and frequent contacts; hence, a transmission network encompasses the social relationships in which these contacts are embedded. Over the years, transmission network studies have made enormous progress in all aspects, including models, methods, and applications. This has granted us a more comprehensive understanding of the spread of HIV and provided effective evidence-based intervention strategies for HIV prevention and control. Additionally, transmission networks have been extensively explored and expanded to examine sexual transmission modes, meanwhile various theories and practical applications of sexual networks have been established and improved. Moreover, with the increasing development of molecular technology, transmission networks have become an innovative research field, enhancing our comprehension of the transmission characteristics of HIV through molecular epidemiology and the high-risk groups of HIV transmission regionally, nationally, and even globally. These networks reflect the characteristics of the HIV epidemic from different perspectives. We herein provide a review of transmission networks, including relevant definitions, applications, limitations, and advances of the studies on HIV transmission networks, covering social, sexual, and molecular networks [Figure 1].

Social networks
Definition
A social network comprises a group of social participants or node members connected by one or more relationship types. Social network analyses include studies of both egocentric (or local) networks and sociocentric (or sociometric, complete, or global) networks. The egocentric model is composed strictly of direct connections to the focal individuals (ego), while the sociocentric network contains direct and indirect connections, as well as...
Characteristics of sexual networks such as size, composition, and density, were found to be associated with HIV risk behaviors such as sharing syringes, drug addiction, having multiple concurrent sexual partners, and practicing unprotected sex.

Structural network characteristics can be described from multiple perspectives, such as centrality and groups. We can also apply network-level measures including size and density, in order to perform a global assessment of the network. The three most common centrality measures are the degree, closeness, and betweenness. Degree refers to the number of links to and from a particular individual, whereas closeness measures the average distance between a node and every other node in the network, and betweenness is the frequency of ties with which an individual is on
the shortest path connecting pairs of others in the network. Groups are defined as any subset of a network. Finally, K-core, n-Clique, and k-plexes are proper parameters enabling to determine the network structure of social networks based on which individuals are in each specific group. Network-level measures apply to an entire network, which provides indexes for the network structure, including size, density, reciprocity/mutuality, clustering, centralization, etc.\textsuperscript{[12]} With these theories and measures, the characteristics of social networks can be described systematically and scientifically to fully explore the connections between social networks and the HIV epidemic.

Applications
Social network analyses can be applied in diverse fields of HIV research, prevention, and treatment. Social networks could be utilized to collect valuable information regarding the HIV epidemic, such as investigating the behaviors of individuals within the network regarding their risk factors for HIV infection and their perceptions of acceptable prevention strategies,\textsuperscript{[13,14]} as well as to gain access to relevant populations and study the patterns of high-risk behaviors including drug use and sexual behaviors.\textsuperscript{[15,16]} Social networks facilitate the collection of static information on the network members and their relationships and also enable to explore their changing characteristics from the perspective of dynamic development.\textsuperscript{[17,18]} Furthermore, social networks provide means of HIV prevention interventions that are low-cost and sustainable, which can be tailored and applied to different populations such as peer education interventions to promote HIV risk reduction among intravenous drug users (IVDU) and to reach men who have sex with men (MSM) for HIV testing and linkage to healthcare.\textsuperscript{[19,20]} Social networks are also extremely useful for optimizing HIV medical care and adherence.\textsuperscript{[21]}

Limitations
Despite all the previously mentioned advantages, the social network approach also has several drawbacks. First, HIV patients may have been infected for several years by the time of diagnosis, and social networks data are usually collected via face-to-face interviews or self-reports; consequently, there may be recall and social desirability bias.\textsuperscript{[22,23]} Second, most index participants are recruited through targeted outreach, which may introduce the possibility for selection bias.\textsuperscript{[24]} Third, linkages may have changed during the investigation period, thus the statistical validity of relationships can be misinterpreted.\textsuperscript{[25]}

Advances
A significant amount of progress has been made in applying social networks to HIV-related fields in recent years, including the use of an online social platform and other methods to improve the ability of social network analysis and intervention, further application of social network analyses to understand a variety of characteristics of the spread of HIV, as well as in-depth exploration of various possibilities for social network AIDS intervention. Furthermore, online tools have been used to boost the power of social network methods for HIV transmission interventions, including social media such as Facebook. Young et al\textsuperscript{[26]} discovered that Facebook can serve as an innovative forum to increase HIV prevention discussions and HIV testing requests among at-risk groups. MacGowan et al\textsuperscript{[27]} recruited MSM online for HIV self-test distribution by mail as well as raised testing and awareness of HIV infection among participants and their social network members, which may be conducive to preventing the transmission of HIV.

Social networks also help us to constantly analyze and understand the characteristics of HIV transmission. Fujimoto et al\textsuperscript{[28]} identified the social mechanisms that drive collective venue avoidance among the network members of younger African-American MSM. The study tested two hypotheses: 1) homophily, with socially connected partners being more likely to avoid the same venues; and 2) popular opinion leaders, such as younger African-American MSM, who are more likely to avoid attending venues. The first hypothesis was supported in both study cities, meanwhile the second hypothesis was supported only in Chicago, suggesting that the structural patterns of venue avoidance differ between cities. Phillips et al\textsuperscript{[29]} examined the impact of social network factors on the use of Pre-exposure prophylaxis (PrEP) among young MSM (YMSM) and young transgender women. They found that out of the 906 study subjects, 7.0% reported using PrEP in the prior 6 months, and recent PrEP use was associated with both individual- and network-level factors, indicating the necessity to raise the attention to social networking factors for HIV prevention activities. Sterrett-Hong et al\textsuperscript{[30]} explored the potential impact of non-parenthal adults (NPAs) in the YMSM social network by sampling among 169 YMSMs from various ethnic groups and their social network data to analyze the correlation between YMSM’s emotional closeness with NPAs and their subsequent substance use. They found that closer relationships with NPAs are associated with a comparatively lower prevalence of substance use after 6 months, including binge-drinking and marijuana use. Messer et al\textsuperscript{[31]} investigated the social network characteristics associated with depressive symptoms and social support among HIV-infected women of color (WOC). Data on the social networks of 87 HIV-positive WOC were collected, and the results revealed that both functional and structural factors of social networks can be a source of support to alleviate depression and assist the affected women in obtaining appropriate care.

Exploring interventions and evaluating their effectiveness are also important applications of social networks. Wong et al\textsuperscript{[32]} found a strong correlation between the chemsex (sexualized drug use) of MSM and HIV infection. However, if the community has access to proper interventions, the propensity of MSM who are medium or intensive drug users may take PrEP to offset their risk of HIV transmission. Besides, mobile applications can reach medium-threshold MSM drug users. Bogart et al\textsuperscript{[33]} conducted a pilot randomized controlled trial designed to promote HIV-positive individuals as advocates in their social networks through a series of intervention strategies.
A total of 99 HIV-infected cases were enrolled in the study, including 51 in the intervention group and 48 in the control group, plus 58 members from their social networks, to assess the changes in cognition, attitudes, and behavior before and at 5 or 8 months after the intervention. The participants’ subsequent positive attitudes and favorable evaluations delineated the high acceptability and feasibility of the intervention. The study concluded that placing HIV-positive individuals at the heart of prevention can help reduce stigma and improve prevention effectiveness across social networks.

**Sexual networks**

**Definition**

Sexual networks consist of people having either direct or indirect sexual contact with each other, usually referring to the retrospective sets of individuals who have engaged in any sort of sexual contact with each other during a specified period of time. Sexual networks that are critical to sustaining the spread of HIV in communities have several key characteristics, including a sufficient number of individuals, a medium to high density of contact, the centrality of infected persons, and sexual partner selection patterns.

Sexual networks can overlap with social networks to some extent, but they also have their own unique focus and characteristics. Sexual networks can be illustrated via random graphs using nodes to represent individuals and edges to represent sexual contact. The numbers of edges adjacent to a particular node indicate the degree, and the sets of node degrees denote the degree of distribution of the population. The propagation of HIV is associated with changes in the degree distribution, which can be targeted to help control the HIV epidemic.

**Applications**

The characteristics of sexual networks, such as size, density, and structure, are closely related to HIV transmission. Helleringer and Kohler conducted an influential study of HIV transmission through sexual networks on Likoma Island in Lake Malawi located in South Africa. They discovered that half of all the sexually active respondents had an extensive and stable sexual network. More than a quarter of women were linked to multiple separate chains of relationships. HIV prevalence is higher in sparse regions than that in tightly connected ones, and the uneven distribution of HIV infection in networks implied that network characteristics were a capital determinant of the dynamics of HIV transmission within the population. Choi et al. found that a large male sexual network size with six or more lifetime male sex partners and a steady stream of male or female sex partners were independently associated with a higher occurrence of unprotected anal or vaginal intercourse.

The sexual network model is also characterized by the concurrency and mixing between different subpopulations, contributing to the spread of HIV. Concurrent sexual partnership was defined as two or more partner-ships with overlapping dates, which increases the likelihood that a person will transmit the infection by ignoring the protective effect of partner sequencing. The risk of HIV transmission was affected by the sexual behaviors of both the individuals and their partners, also known as individual concurrency and partner concurrency, respectively. Kenyon et al. performed an analysis of a nationally representative sample of 15,031 persons and found that partner concurrency may increase the risk of HIV infection among black South African women, especially the young.

Sexual mixing patterns play a critical role in HIV transmission. Assortative mixing refers to sexual partnerships among people with similar levels of infection risks. Whilst disassortative mixing occurs when partnerships form between higher and lower risk individuals. Mimiaga et al. investigated black MSM sexual mixing patterns and found that their sexual risk behaviors differ across partner types and the characteristics of their sexual networks, putting this subpopulation at particularly high risk for HIV infection and transmission.

Prevention strategies can be developed according to the characteristics of sexual transmission. By analyzing the data from ten cross-sectional surveys conducted between 2000 and 2013, Aghaizu et al. noted that a combination of behavioral and biological interventions for community-wide risk reduction was essential to progress towards HIV eradication. Rodger et al. conducted a prospective study to follow up 1166 HIV serodiscordant couples with 1238 couple-years. The results revealed that no documented case of HIV transmission occurred within serodiscordant heterosexual and MSM couples who used suppressive antiretroviral therapy (ART) with reports of condomless sex during the study period. Similarly, a large PrEP demonstration project in the United States by Liu et al. demonstrated that the incidence of HIV infection was extremely low despite the high incidence of sexually transmitted infections (STIs). Compliance was higher among participants who reported riskier behaviors.

**Limitations**

The limitations faced during sexual network analyses were similar to those encountered in social networks analyses. First of all, individuals may not report all of their sexual partners due to recall bias. Second, network members from outside the research area normally cannot be enrolled, which may limit the ability to generalize findings for a broader area. Third, there is a potential for distortion of self-reported risk behaviors and sex network members, due to social desirability. Moreover, quantifying the degree in sexual networks using self-reported cross-sectional data on the day of the survey may lead to bias mainly due to the uncertainty of future sexual behavior.

**Advances**

In recent years, research on sexual networks principally focused on acquiring an in-depth understanding of the characteristics of HIV transmission and the formulation of effective and efficient prevention strategies. Liu et al. demonstrated that the incidence of HIV infection was extremely low despite the high incidence of sexually transmitted infections (STIs). Compliance was higher among participants who reported riskier behaviors.
explored the characteristics of female sex workers' (FSWs) sexual networks, and also how they and their sex partners could serve as ‘bridges’ in the spread of HIV/AIDS. The experimental results revealed that FSWs and their commercial sex partners served as key nodes within high-risk groups and bridges between high-risk groups and the general population. FSWs and their clients can act as effective bridges for HIV/AIDS transmission from high-risk groups to the general population by establishing a link between commercial and non-commercial sexual networks. McCree et al. used sexual network analysis to explore the possible reasons for the large decline of HIV diagnoses among Black/African-Americans. The results suggested that there are temporal trends in factors that increase HIV infection and transmission capacity in sexual networks, including the diminishing rates of other sexually transmitted infections among African-American women, reduction of HIV diagnoses among potential male partners, and increased application of viral suppression therapy among HIV-positive heterosexual and bisexual potential male partners.

Xiao et al. demonstrated that HIV self-testing (HIVST) among MSM in China based on sexual networks is feasible, and is a promising strategy for improving the effectiveness of HIVST programs. This approach should be complemented by interventional components that enhance the effectiveness of HIV testing. Okafor et al. determined whether disclosure of HIV status is significantly associated with reported HIV sexual risk behaviors among HIV-positive African-American MSM in six cities in the United States. They found that disclosure of HIV status was common, and was associated with lower HIV sexual risk behaviors. This could serve as a potential guide helping researchers to develop prevention protocols that aim to increase HIV status disclosure within the networks.

Molecular transmission networks

Definition

The high evolution rate of HIV allows us to re-establish connections between viral strains from different individuals. The degree of each individual in the network is defined as the number of links with other individuals. Clusters are defined as connected components of the network containing two or more nodes. Epidemiologic contact information is not necessary for clustering of the molecular network, since the link between individuals does not imply direct transmission, but rather two recently related viral strains, normally referred to as putative transmission chains.

Genetic clustering methods can be roughly divided into two categories: Methods for directly clustering sequence variations using pairwise distance measures, and methods for interpreting variations in phylogenetic subtrees. A combination of both approaches can also be utilized for network establishment. Pairwise genetic distances could be computed with the most commonly used model, the Tamura-Nei 93 (TN93) nucleotide substitution model, and the distance could be estimated directly from the counts of nucleotide pairs in the aligned sequence. On this basis, the distance cut-off needs to be defined for sequences belonging to a certain cluster based on the objectives of the study. Generally, the cut-off value would be lower than 1.5%, but if the purpose of the study was to obtain the most accurate identification of transmission networks and to assess ways through which voluntary contact tracing can detect other individuals who fit these networks, a more conservative genetic distance could be chosen; for example, < 1%. In this case, the spread cluster would be defined as pol sequence similarity of over 99% between any two individuals. With a conservative cutoff, participants are less likely to be included in an incorrect transmission network, but they would also be more likely to be excluded from the cluster to which they belong. Molecular phylogeny is a tree-based model of the mechanism through which gene sequences are linked by common ancestors. Due to host-specific selection, the genetic composition of rapidly evolving viral strains may differ considerably on the time scale of transmission. Poon et al. conducted a population-based phylogenetic analysis in British Columbia, Canada. They generated 100 bootstrap samples by resampling columns from the alignment with random replacements in order to control for uncertainty in phylogenetic reconstruction. For each bootstrap, FastTree2 was applied in the inference of maximum-likelihood phylogenies for huge alignments. To be more specific, all pairs of tips were used with: 1) Patristic distance < 0.02; 2) the tips representing HIV sequences from different infections; and 3) at least one of the sequences derived from that particular individual’s earliest available sample. All pairs of tips that met these criteria in > 50% of the bootstrap trees were used to construct a graph. Other clustering methods were also spotted including the Gap Procedure, Bootstrap and branch-lengths, Cluster Picker, and PhyloPart. Each of these methods has its own advantages and disadvantages, and a consensus is yet to be reached for phylogenetic methods, sampling ranges, gene targets, and/or minimum fragment sizes.

Applications

Molecular transmission networks can enable proper scrutiny of network characteristics on a much larger scale and longer study period than common social or sexual networks. Furthermore, molecular transmission network analyses could be conducted on a national or even global scale. Wertheim et al. conducted a network analysis using all publicly available HIV polymerase sequences in the global database, and the transmission network comprehensively depicted the spread of HIV-1 at the national and international levels. The study reinstated well-characterized transmission clusters and extended other clusters across international borders, although they had been thought to be contained within the national level, and discovered previously undescribed transmission clusters. The transmission network represented all known modes of HIV transmission and reflected the need to consider the global diversity of HIV. Paraskevis et al. applied maximum likelihood inferences on partial pol sequences from 8955 HIV-positive individuals linked by demographic and clinical data. Molecular transmission clusters (MTCs) were defined using two different criteria:
1) Bootstrap support was >75%, and 2) the ratio of sequences from a specific region was >75% compared to the total number of sequences within the network. They found that 3700 (41%) sequences belonged to MTCs, and the proportion exhibited remarkable discrepancies between countries and subtypes. Finally, the demographic and clinical factors associated with MTC clustering were assessed.

Moreover, molecular transmission networks have been applied to various other fields such as describing the transmission dynamics of HIV,[67] determining the prevalence and features of transmitted drug resistance (TDR),[68] exploring the epidemic history in a given area,[69] identifying individuals belonging to a highly relevant transmission group to implement selective and targeted preventive interventions,[61] detecting a recent outbreak of a series of transmission events,[70] and investigating hot spots of rapid transmission.[71] As HIV treatment centers continue to accumulate HIV sequence-based genotypes, there is a growing interest in implementing this resource to provide information on real-time HIV prevention and control measures.[34,71]

**Limitations**

As any other scientific model, HIV molecular transmission network analyses contain a certain number of limitations. First, since phylogenetic methods are unable to confirm direct transmission or the order of transmission, conclusions on transmission patterns are drawn based on indirect connections. Second, the contributions of undiagnosed individuals, diagnosed individuals without genotyping, and individuals tested before or after the study period could not be assessed.[72] Third, HIV molecular epidemiology hypothesized that outbreaks of HIV transmission (whether spatial, temporal, or both) are usually reflected as a cluster, but the reverse is not always true; in other words, clusters do not always reflect outbreaks.[35]

**Advances**

Over the past decade or so, research projects of molecular transmission networks have primarily focused on identifying transmission networks among populations,[73] monitoring HIV transmission dynamics,[74] identifying long-term trends,[75] establishing innovative analytical methods,[76] developing intervention strategies,[77] and evaluating the effectiveness of interventions.[78] Zhang et al.[73] used partial pol gene sequences from 486 HIV cases within 15 districts of Anhui Province to construct transmission networks. Four transmission subnetworks of high-risk individuals (degree ≥ 4) were detected, including two CRF01_AE transmission subnetworks comprised of those with average ages of 67.9 and 61.5 years old. Infection was most likely transmitted through heterosexual contact, while the other two CRF07_BC transmission subnetworks were mainly spread by MSMs, with average ages of 31.73 and 34.15 years old. Nduva et al.[74] used 658 sequences to investigate HIV-1 transmission networks among MSM, IVDUs, FSWs, and heterosexuals in coastal Kenya. They found that most clusters (85%) consisted of sequences from the same risk group, delineating frequent

**Combinations of networks**

There is no panacea for HIV prevention, and only a combination of prevention strategies can curb the HIV epidemic.[79] Likewise, there is no universal method for understanding the characteristics of HIV propagation. A comprehensive understanding of HIV transmission can only be achieved by combining multiple network methods. Transmission networks, including social, sexual, and molecular transmission, are of immense value for preventing and controlling the HIV epidemic. Each network has its own focus and advantages, as well as its own limitations. Effective combinations of various transmission network methods can confer a more comprehensive and multi-dimensional picture of the HIV epidemic, which is conducive to a well-rounded, efficient, and accurate description of the characteristics of HIV transmission networks and the formulation of powerful prevention and intervention measures.

Phillips et al.[29] investigated the role of social and sexual network factors in the use of PrEP among YMSM and transgender women in Chicago. They used the data gathered from participants within a longitudinal cohort study, and adopted both social and sexual network models to analyze the individual- and network-level factors associated with PrEP utilization in the prior 6 months, which identified significant correlations at both levels. Using social and sexual network data, Imahashi et al.[80] explored the impact of HIV status disclosure to social and sexual partners among YMSM. They noticed that
respondents who had overlapping sexual and social relationships with their partners were less likely to be unaware of their partners’ HIV status, which highlighted an association between overlapping relationships and partners’ knowledge of HIV status among YMSM.

Withal, determining ways to best combine molecular and social network data is crucial since both approaches provide necessary but insufficient information on the infection’s transmission route. Molecular transmission network technical analysis can be applied to support or refute the validity of epidemiological HIV transmission networks. Wertheim et al extrapolated and compared sexual/drug-use partners and genetic transmission networks using HIV sequences and partner naming data from 1342 HIV-infected individuals in New York City between 2006 and 2012. Genetic linkages provide more reliable evidence for identifying potential transmission partners than partner naming, which demonstrates the importance and complementarity of epidemiological and molecular genetic data in characterizing regional HIV epidemics. Tookes et al conducted an epidemiologic investigation of ten anonymously newly identified HIV seroconversions, out of which seven acute HIV seroconversions were found to be associated with treatment and achieving viral suppression, among which six cases were epidemiologically and/or socially associated with at least two other seroconversions. Analysis of HIV genotypes showed that two individuals had molecular connections at a genetic distance of 0.5%. This risk network identified complex transmission dynamics, which could not be explained exclusively by epidemiological approaches or molecular analysis.

Combining various network data could also provide valuable information for prevention and intervention. Kwan et al integrated the phylogenetic, clinical, and behavioral data of newly diagnosed HIV patients to analyze transmission dynamics using an information diffusion model. In this case, HIV sequences were used to construct the genetic network, and the transmission cascade was extracted from the network built. The genetic associations of MSM with different sexual network channels and risky behaviors also varied. According to the usage patterns of mobile applications, MSM differed significantly in network metrics, which implies that social networks had influenced the transmission networks. With continuous longitudinal data collection, the applied model can enhance HIV molecular epidemiological surveillance to inform future intervention plans. Little et al used HIV-1 pol sequence data from recently infected persons as well as their social and sexual contacts to design a highly connected transmission network. They also developed a transmission network score (TNS) to assess the risk of HIV transmission from a newly diagnosed individual to a new partner and targeted preventive measures. They found that the TNS was highly correlated with transmission risk behaviors and outcomes, hence it could be implemented to tailor effective prevention interventions for populations at a higher risk of HIV transmission.

In a nutshell, this study introduced the basic definitions, common applications, existing limitations, and recent research progress overviews of social, sexual, and molecular transmission networks, as well as the joint application of those networks in related research. The advantages provided by combining various transmission networks, including but not limited to the comprehensive-ness of social networks, depth of sexual networks, and precision of molecular transmission networks, will undoubtedly increase our understanding of HIV transmission networks; the nature and dynamics of HIV transmission networks are essential for monitoring, preventing and ultimately, eradicating HIV.

Conflicts of interest

None.

References

32. Wong NS, Kwan TH, Lee KCK, Lau JYC, Lee SS. Delineation of

30. Sterrett-Hong EM, Birkett M, Kuhns L, Zhang D, Mustanski B. The

26. Young SD, Jaganath D. Online social networking for HIV education

34. Amirkhanian YA. Social networks, sexual networks and HIV risk in


17. Costenbader EC, Astone NM, Latkin CA. The dynamics of injection

28. Fujimoto K, Wang P, Li DH, Kuhns LM, Amith M, Schneider JA. The impact of closeness to non-parental adults in social networks on


