Personal Risk Taking and the Spread of Disease: Beyond Core Groups

Richard B. Rothenberg, John J. Potterat, and Donald E. Woodhouse

Disease control efforts directed at human immunodeficiency virus are predicated on the need to reduce personal risk behaviors; that approach may not adequately reflect the complicated interplay between personal behaviors and the social setting in which they occur. Efforts to date, including the application of population ecology, the development of the core group hypothesis, and the use of compartment models to describe disease transmission, have aided in understanding the dynamics of transmission and have highlighted the relationship between personal risk taking and population risk. An area for further development is the application of the techniques of social network analysis to infectious disease spread. Initial work suggests that social structure may act as a barrier (or facilitator) in disease transmission and that the epidemiologic impact of a risky act varies with the social setting. The local context for risk behaviors has important implications for the dynamics of transmission.

Many of the traditional personal sexually transmitted disease (STD) risk factors—a large number of sex partners, frequent partner change, casual encounters—may be viewed as surrogates for the probability of encountering an infected partner [1]. Thus, an underlying principle in STD control has been that "...change of the patient's sexual behavior is...an essential element in the prevention and control of STD" [2]. The HIV control program in the United States is based on such an axiom: "Risk behaviors for HIV transmission are clearly identified. Reduction of risk behaviors results in reduction of HIV incidence. Public health interventions (individual or community) can result in reduction of risk behaviors" [3].

A number of studies provide some support for this optimistic strategy [4], and a cost-benefit analysis based on assumptions of only modest efficacy point to the value of offering counseling services to human immunodeficiency virus (HIV)-infected persons [5]. On the other hand, skepticism about the value of behavioral interventions has also been expressed [6]. Debate on the relative merits of behavioral approaches obscures a more fundamental question: Are the underlying assumptions about risk behavior and its impact on HIV transmission valid? Put another way, do the premises on which this strategy rests accurately reflect HIV transmission dynamics?

At face value, the logic seems sound: Since HIV is passed from person to person during risky acts, minimizing such acts should minimize transmission. This notion of transmission presupposes that risks are like votes, and their simple total determines the outcome of the election (epidemic). Several concepts suggest that this might not be the case.

First, all human actions take place within a social context. Epidemics are not just the consequence of many risky acts but the result of a complex interaction of behaviors that may be sequential or parallel, contemporaneous or asynchronous, contiguous or dispersed. In short, we must consider the intricacies of human interactions, not just the aggregate effect of unrelated acts.

Second, this complex interaction occurs within definable groups. Our traditional focus on individual acts has led to characterization of groups based on specific behaviors (homosexual men) or on some overall assessment of risk (high-risk). Such categories may be useful for descriptive purposes but fail to capture important attributes that must be considered to understand disease transmission. Precise description of groups using demographic, behavioral, spatial, and temporal characteristics offers greater potential insight. Two examples provide illustrations of how these characteristics might be defined: persons aged 15-24 years who injected drugs and who resided in South Central Bronx, New York, from 1980 to 1985; a group of 25- to 44-year-old upper middle class, monogamous, homosexual men, living on New York’s Upper West Side in the early 1990s, who are socially and occupationally, but not sexually, connected.

Third, all groups, regardless of their attributes, are part of a larger society and are subject to myriad social forces. These forces act as effect modifiers or cofactors to enhance or diminish the actual influence of any risky act on the spread of disease. For example, high-risk transmission behavior may be modified by other risks: a young man who injects drugs, shares syringes, and never uses condoms may be the victim of homicide before he contracts or transmits disease. In addition, the overall prevalence of HIV in a community can influence transmission in unpredictable ways, depending on the distribution of infected persons. A higher-prevalence area with homogeneous distribution may be subject to less transmission than a lower-prevalence area with a highly focal pattern.

In recent years, considerable effort has been devoted to exploring how individual acts, group membership, and social forces interrelate. The resulting paradigm challenges the notion that epidemics are simply the sum of unrelated individual acts.
Rather, it replaces individuals with social groups as the foci for disease propagation. The paradigm is characterized by a plethora of terms whose meaning and interrelationships are elusive: personal behaviors, risk groups, reproductive rate, compartments, core groups, contact mixing patterns, social networks, hierarchical and spatial diffusion, and others. A consideration of some of these terms, and recent work devoted to their elucidation, may help in establishing connections.

Ecological Concepts

The notion of sexual spread of an infectious agent is captured in a now-familiar concept borrowed from population ecology [7]: The reproductive rate of a disease \( R_0 \), that is, the number of new infections generated by a given infection, is directly related to the transmissibility of the disease \( (\beta) \), the effective rate of partner change \( (c) \), and the duration of infectiousness \( (D) \): \( R_0 = \beta \times c \times D \). Whether or not an epidemic propagates thus depends on some intrinsic properties of the parasite-host interaction as well as on the behavioral characteristics of the host. An important corollary to this basic formula is that the effective rate of partner change is equal to the mean rate of partner change in the population \( (m) \) plus the variance of the rate of partner change divided by the mean: \( c = m + \sigma^2/m \). The most sexually active people thus play a disproportionate role in transmission. Recent work suggests that the effective rate of partner change can be estimated directly from the actual rate of partner change in a sexually active population such as STD clinic attenders [7].

Core Groups and Compartment Models

The use of population ecology to describe disease propagation developed in parallel with the notion that the interaction of groups at differing risk can be used to describe transmission. Hethcote, Yorke, and Nold [8, 9] made the fundamental observation that the distribution of gonorrhea was nonrandom and, in fact, was dominated by areas with intense transmission. They noted that the overall \( R_0 \), as derived from contact tracing data, was \(<1.0\), despite the dramatic increase in gonorrhea incidence in the early 1970s, and concluded that the true \( R_0 \) must be a weighted average of those who transmit and those who do not. They further postulated that the latter were probably aggregated in small, distinct groups—so-called core groups. Such groups might be responsible, directly or indirectly, for most gonorrhea transmission. They described a system of differential equations that could be used to solve for the equilibrium state of a gonorrhea epidemic. Their system involved the interaction of eight groups (or compartments) with different characteristics, using arbitrary estimates for transmission parameters but incorporating the notion of core groups of active transmitters (i.e., some of the compartments had core characteristics).

The use of core groups in modeling transmission has since been extensive and often innovative. For example, Whitaker and Renton [10] examined the upturn in gonorrhea among homosexual men in the early 1990s and attributed it to a decline in sexual activity among those who were not in core groups. They argued mathematically that noncore groups were adhering to risk reduction guidelines, thereby reducing the sex partners available to those in core groups and intensifying transmission within such groups.

The basic notion of compartments with varying risk, derived in part from classic epidemic theory [11], was further developed by others [12-14] with a particular focus on HIV transmission. However, Gail and Brookmeyer [15] point out that such compartment models are useful as a simulation environment for studying disease transmission but are not useful for prediction. Anderson and May [16], summarizing estimates of transmissibility and partner change, highlighted an additional limitation to the method: the lack of empirical data for parameter estimation. Particularly for situations wherein numerous groups are examined, the number of parameters to be estimated quickly becomes unmanageable. For example, if groups are defined by age, sex, race, location, and risk, and there are only two levels of each variable, then parameters must be estimated for the interaction among 32 compartments.

That groups interact in an epidemiologically meaningful way has nonetheless been supported by some empirical evidence. In a study of gonorrhea transmission in Colorado Springs, data-based correlates for duration of infection and for partner mixing patterns were described [17]. The force of infectivity, defined as the number of potentially infectious days that persons with gonorrhea generate in their contacts, was used as a surrogate for duration of infection (this term should be distinguished from the force of infection [6, 14], a true hazard rate). The force of infectivity was developed to provide an empirical measure, derivable from routinely collected program data, that could approximate the period of infectiousness to which others were exposed. Self-selection, or the frequency with which persons from a defined subgroup had contact with others in their subgroup (as opposed to contact with those outside their subgroup), was a surrogate for mixing patterns.

Analysis of these contact tracing data suggested that persons with a majority of contacts outside their subgroup generated a force of infectivity that was four times greater than those who remained within their subgroup. These observations are in keeping with theoretical work showing that assortative (within-group) mixing led to quicker initial epidemic growth while disassortative (between-group) mixing generated a larger epidemic in the long term [18, 19]. A comprehensive analysis by Garnett and Anderson [19] demonstrated that the assortative mixing pattern was predominant in Colorado Springs [20] and New York State [21] and served to maintain gonorrhea endemicy in small areas of intense transmission, the aforementioned core groups. An important finding implicit in their work is that communities occupy a spectrum of mixing patterns and coreness.

Further development of the core group concept has been constrained by its popularity. A plethora of definitions, enumer-
ated elsewhere in this issue [22], has emerged. Perhaps it should be recognized that the original working concept (that core groups were socially, behaviorally, and geographically stable aggregations of people whose sexual activity permitted continuing passage of organisms among them [8, 9]) served Yorke, Hethcote, and Nold's analytical framework well but did not provide an axiom. The proliferation of definitions may simply reflect that analytical enterprise: Core groups work well as a unitary compartment whose contents remain unspecified, a temple of the unfamiliar. Rather than seeking new operational definitions, some investigators have focused on the social, behavioral, and geographic nature of these groups and, by examining the forces that affect groups of people, have moved our understanding in a different direction.

Risk Behaviors in Risk Space

Let us combine two previous examples: the young man who used injectable drugs and never used condoms, and his community in the South Central Bronx. The social forces acting on that man, within that group, have been described in great detail by Wallace [23–25] and Wallace and Fullilove [26] and summarized by others [27]. Briefly, these investigators examined the relationship between the withdrawal of municipal services, notably fire extinguishment, and the loss of population from the South Central Bronx through migration to nearby and in some cases more distant health areas in the 1970s. With these population shifts came a measurable worsening of social conditions (poverty, overcrowding) and a parallel shift in the concentration of social and medical problems in the Bronx (tuberculosis, high-risk pregnancy, cirrhosis mortality, overcrowding, violent deaths, and drug-related deaths). This population displacement and community disintegration are coterminal with the considerable spread of HIV that occurred in the Bronx in the 1980s. These investigators conclude that the social events were instrumental in propagating HIV transmission through geographic displacement, loss of traditional social ties, and the accompanying breakdown of barriers to risk behaviors. Consider again the young man in this community context: it is clear that the social setting would facilitate transmission that might result from his activity.

The implication is that the national HIV epidemic is the weighted average of events among interconnected groups of people who are subject to a spectrum of social forces. Understanding the dynamics of HIV transmission, then, is contingent on understanding the interaction of persons within groups within a larger social context—risk behaviors in a risk space. Human interactions can be defined in three dimensions, at least two of which have a well-defined metric for distance: time and geographic space. The third dimension, social connection, is far more difficult to measure. Nonetheless, the hypothetical construct of risk space has intuitive appeal. All three elements are necessary for an infectious agent to be transmitted: The requisite behaviors for HIV transmission, in the absence of an appropriate temporal, geographic, and social setting (i.e., the right network), may not result in disease propagation.

This construct was succinctly summarized by Catania et al. [28]: “To make a more exact assessment of the risk associated with a person’s sexual activity, one would need to know the size and characteristics of his or her sexual network, the prevalence of HIV infection across social strata in which those sexual networks are embedded, the type and frequency of sexual practices engaged in, as well as information on donor infectivity and host susceptibility.” In this context, the reproductive rate, core groups, and compartment models address part of a more complex system.

Social Network Analysis

Perhaps in an attempt to understand risk space, considerable interest is now focused on the social networks of persons who may be at risk for HIV infection. The materials and methods of social network analysis are well established [29] but have only recently been applied to the STD and intravenous drug use context. Largely through the sponsoring efforts of the National Institute on Drug Abuse and the Centers for Disease Control and Prevention, a number of projects have been begun that examine the role of social networks in disease transmission.

Social network analysis has a basis in graph theory [30]. Persons (often called “egos” or “stars”) constitute nodes or vertices in a graph. The vertices are connected by lines (edges) that may be singular, multiple, cyclic, or noncyclic and may have a number of properties that describe the connection. Such connections are usually defined at the time of data collection. In many social network studies, these paths may imply friendship, kinship, a work relationship, a neighborhood association, or simply knowledge of another person. In disease transmission studies, the path can imply a route of transmission.

If there is a path of some length between each person and every other person in the network, the group is said to be connected and each person in the network is reachable. A portion of the graph (a subgraph) may have these properties as well. Numerous measures have been devised to characterize the network as a whole and to characterize the position of persons within the network. For example, a network is said to be dense if many or all of the available connections among people in the network actually exist, as in a kinship network. In such a situation, there are many paths from one person to any other person. Network boundaries may be thought of as clearly defined or fuzzy, depending on the ease with which a person can enter the group or some preexisting size criterion. For example, a top corporate management group or a professional athletic team would have clearly defined boundaries. A group of intravenous drug users in a shooting gallery might constitute a network whose boundaries are constantly shifting.

Persons within a network may occupy a central or peripheral position. Freeman [31] proposed three general classes of centrality measures: distance, usually defined as the mean or me-
median path length from a person to all others in the network; adjacency, the number of people that person is directly connected to; and betweenness, the extent to which a person stands in the path between other people. In many social network analyses, these measures—all of which reflect the relative position of an individual with regard to others—have been used to assess a person's prominence or influence. Each captures a slightly different aspect of prominence: distance reflects how close a person is to all others in the network; adjacency, the opportunity for direct connection to others; and betweenness, the opportunity for a person to act as a barrier or conduit for information transfer. Such connections have direct meaning when network analysis is applied to infectious disease transmission. In the case of HIV, the connection is sexual or needle-sharing or both. The social metric and centrality measures using the metric are directly related to the probability of transmission. The intensity of risk-taking activity and the use of personal preventive measures can provide a weighting system for assessing centrality within a network and for evaluating the overall threat that a network structure poses for its members.

Several study sites are now pursuing the use of social network analysis to gauge HIV transmission, and early results encourage continued exploration. For example, in a network of 595 persons at presumed high risk for HIV in Colorado Springs (an area of low endemicity with little or no endogenous transmission [32]), almost 6000 persons were identified as part of the overall network of the original respondents [33]. Network techniques for analysis [34] demonstrated that at-risk persons could be grouped into a number of connected components (a group of people within which a pathway exists of some length between any two persons; such a grouping does not imply that each person knows everyone else directly). In this structure, which uses sexual contact as the relationship of interest, there was a single large connected component (965 persons) and a number of much smaller connected components. The majority of HIV-infected persons were found in the smaller components, suggesting that they were socially isolated from the main body of persons in the network (figure 1). Thus, despite considerable risk-taking behavior that was well-documented in the group, propagation has not occurred, and the social network structure appears to act as a barrier to transmission. The dynamic nature of networks underscores the need to pursue such hypotheses in greater depth.

In a similar study in Flagstaff, Arizona, numerous small networks of injecting drug users were described [35]. The prevalence of HIV infection was low and confined to several small networks. By enumerating the nature of contacts between individuals within a group and the points of intersection of the
groups, the investigators could develop a predictive model of HIV spread (predictions they hope to belie with active intervention that is part of the project).

In a more theoretical vein, Wallace and Wallace [36] have combined classic epidemic theory with notions of social network interaction to show that the spread of HIV is a function of the height of the epidemic in an epicenter and the strength of the commuting field between that center and some outlying area. The term "commuting field" is a literal one that refers to the frequency with which persons commute from one area to another. It is used as a surrogate for the amount of contact involving risk behavior that might take place between two areas. Such a construct could be used to assess the potential for spread on the basis of empirical data obtained from the actual networks involved. It is a large-scale correlate of the predictive model actually delineated by Trotter [35] in Flagstaff.

Each of these observations contains several notions that may be important for disease control, though such notions are clearly speculative at this early stage of development. First, segmentation of social networks may be an important method for interrupting transmission. In practical terms, closing of bathhouses and shooting galleries may represent such segmentation. For an area such as the South Central Bronx, restoration of municipal services and revival of an economic base may serve to resegment what are now diffuse, amorphous, and unpredictable networks. Second, social network methods may help to identify persons whose position in the network gives them influence beyond their personal behaviors. The efficacy of educational and other services could be amplified considerably by addressing them to people with important social roles.

Human beings copulate in a context. The interaction of people within networks provides a link in our understanding of the relationship between individual behavior and transmission dynamics. Considerable work remains to delineate the path from reduction of risky behavior to the reduction of disease transmission, but one conclusion from this construct is that generalized, untargeted methods may be of less value than an approach that focuses its resources on important group structures. Perhaps we can abandon the notion of risks as votes, but still paraphrase the late Thomas P. "Tip" O'Neill: all disease control is local.

References


34. ProGamma, Inc. GRADAP, graph definition and analysis package. Version 2.1. Amsterdam: Inter University Project Group, University of Amsterdam, 1989.