Effect of Immune Response on Transmission Dynamics for Sexually Transmitted Infections

Suvendrini Lena, Babak Pourbohloul, and Robert C. Brunham
University of British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada

Transmission dynamics for sexually transmitted infections (STIs) exhibit a large degree of heterogeneity, much of which has been attributed to behavioral variability. However, because STI transmission concentrates among individuals who frequently change sex partners, immune responses also are likely to contribute to the heterogeneity in STI transmission dynamics. We review both theoretical and experimental data on the effects of immunity on STI transmission dynamics. We conclude that research should be directed more intensively toward the characterization of sexual network structures, together with qualitative and quantitative analyses of the immune responses of individuals who are identifiable within the network structure itself. Elucidation of the immunobiological and behavioral factors that shape STI transmission should inform better STI prevention and control programs.

Control of sexually transmitted infections (STIs) has been a tremendous public health achievement. However, by most measures, this success has been incomplete. Questions remain with regard to why control of STIs is not even more successful; whether interventions for the control of STIs can be delivered in a more cost-effective way; whether selected STIs can be successfully eliminated from human communities; and whether vaccines can be developed and effectively used to prevent the major STIs. These and other questions drive much of the current agenda in STI research.

Recent developments in theoretical frameworks and empirical data on the immunoepidemiology of STIs offer hope that answers to some of these questions can be obtained. In this article, we review insights emerging from the mathematical analysis of models of STI transmission and selected data on immunoepidemiological features of STI transmission. We use these data to explore and explain, in part, the large degree of heterogeneity in STI transmission dynamics that is not attributable to behavioral variability. In particular, we describe how different patterns of immune selection within “core groups” may act to structure populations of STI-causing pathogens, influence pathogen virulence, and produce variability over time in the probability of transmission. Viewing STIs from the point of view of their ecological success within core groups can explain characteristics of STIs, such as evolution toward and away from virulence and extreme antigenic variability, and phenomena such as “original antigenic sin.”

MATHEMATICAL ASPECTS OF TRANSMISSION VARIABILITY

Mathematical modeling of STI transmission has progressed steadily over the past 3 decades; in fact, STIs were among the first group of infectious diseases to be studied in detail via mathematical models [1, 2]. One of the important results gained from this type of analysis was the identification of a threshold value, the basic reproductive number ($R_0$), which determines the spread of a specific pathogen within a susceptible population after the introduction of a few infected individuals [3]. Simple models suggest that $R_0$ is proportional to the average probability of transmission per partnership ($\beta$), the average duration of infectiousness ($D$), and the average number of sexual partnerships formed per unit of time ($c$): thus, $R_0 = \beta c D$. These simple parameters...
belie the complexities of STI transmission, including the effects of immunity. For instance, the probability of transmission ($\beta$) between an infected and a susceptible individual is the outcome of a complex relationship between the genetic properties of pathogens and hosts and the immune response. In addition, immune responses determine the duration of infectiousness ($D$).

Most mathematical models of disease propagation are based on the assumption that populations are fully mixed, meaning that an infected individual is equally likely to spread infection to other members of a population. In these models, the population is divided into smaller subgroups that correspond to different epidemiologically plausible compartments. The assumption of a fully mixed model is necessary to enable the creation of a set of differential equations describing the movement of people from one compartment to another. The simple situation in which recovery from infection does not confer immunity can be represented by 2 classes of individuals, susceptible (S) and infected (I), when individuals return to the susceptible class after clearing infection (i.e., SIS models). The SIS approach is often used to model STI transmission dynamics but entirely ignores the effects of immunity. For infectious agents that engender immunity, another compartment, the recovered (or immune) class (R), is needed (i.e., SIR models).

These simple models have been studied extensively over the past 3 decades, resulting in very interesting analytical conclusions [3]. However, as infection propagates through real heterogeneous human populations, a spectrum of clinical trajectories are produced that cannot be represented by the SIS and SIR models. Although more-complex compartmental models have been developed, they deliver epidemiological realism at the expense of analytical tractability [4, 5].

Empirical analyses of patterns of human sexual contact are not consistent with the assumption of a fully mixed model. In the real world, marked variation occurs in the number of sexual contacts per person per unit of time when people do not mix randomly, and each individual has contact with only a small fraction of the entire population. Such data typically generate a nonnormal distributional relationship between rates of partner change and their frequency within the population (figure 1A) [7]. Network models represent an alternative approach to modeling this heterogeneity in patterns of sexual contact. Like compartmental models, network models also need to contend with immunological heterogeneity within the network and over time. For a network model, the question can be formulated as follows: how can we assign different probabilities of transmission between any 2 individuals, on the basis of their immunological status and their location within a sexual network? It is worth considering that immunological variability might not be randomly distributed throughout a network and that immunological differences might be dependent on the network topology itself.

![Figure 1](image-url)

**Figure 1.** A, Distribution of the no. of new sex partners in the prior 12 months in a random sample of 1481 individuals in the United States (data are from [6]). B–D, Graphical representations of 3 different types of networks: B, regular lattice network; C, Poisson random graph; and D, scale-free network. In each panel, a 2-dimensional geometric configuration of each type (right) and the probability distribution of the no. of connections ($K$) for a node in each configuration (arbitrary scale; left) are shown. In panel B, each individual (vertex) has 4 neighbors, except for those who are on the edges of the square. If the size of the network is very large, the role of the vertices on the boundaries can be ignored. In the Poisson random graph (C), the no. of contacts is more or less uniformly distributed around a specific value (the peak of the curve in the left panel of C), thereby excluding highly connected individuals, who play the role of core-group members within sexual networks. In the scale-free network (D), the variability in the distribution of the no. of contacts is shown (left). The Y axes in panels C and D can be interpreted as the probability that a person has $K$ contacts, and the difference in connectedness specific to these types of networks is shown (right).
The analysis of a system from the point of view of network theory can be undertaken at the following 3 levels: the organizational structure of networks, the effect of network structure on disease transmission, and the variation in network structure over time [8, 9]. The application of network theory to STIs is relatively new, and all 3 levels have not been studied to the same degree, perhaps owing to the infancy of existing theories.

The first and second levels of network analysis are concerned with describing the structure of a network and how network structure determines disease transmission. The influence of the network structure on the dynamics of disease propagation has been explored at the theoretical level by use of percolation theory [10–12]. Percolation theory is a branch of mathematical physics that originally was developed to study how liquids move through an aquifer or how fluid molecules spread through porous media and is now being used to study how diseases spread through a population. Percolation theory studies the formation and structure of clusters of cases of infections, on the basis of contacts between individuals.

Although most diffusion processes of practical interest, including the spread of STIs, take place in complex networks, the bulk of diffusion studies have focused on model systems such as regular lattice or Poisson random networks (figure 1B and 1C). Both regular and Poisson random networks have a characteristic scale of connection for each node in the network. Regular networks such as a square lattice (figure 1B) are structures in which every individual (vertex) has the same number of contacts. Although this assumption simplifies the mathematical analysis of the network, it does not realistically represent the web of human sexual contacts. Poisson random graphs are network structures in which the link representing the contact between any 2 individuals can be created with a fixed probability for all links. Although this type of network overcomes the limitation of regular lattice networks by incorporating different numbers of contacts for different individuals, the distribution of the number of contacts is centered around an average value. Therefore, despite differences, individuals have more or less the same number of neighbors; thus, this structure also fails to capture the uneven heterogeneity that exists in human sexual networks. In fact, a recent groundbreaking study showed that the social networks responsible for the spread of STIs exhibit a scale-free structure, of which the topology fundamentally deviates from the topology of regular lattice networks and random networks (figure 1D shows a graphical representation of this type of network structure) [13].

Scale-free networks exhibit a power-law decay \( k^{-\gamma} \) of the cumulative distribution \( P(k) \), where \( P \) is the cumulative probability distribution, \( k \) is the number of links (partners) connected to each node (i.e., sexually active person) in the sexual network, and \( \gamma \) is a characteristic value derived from the network structure [14]. The implication of this type of network structure is that a minority of nodes, or people, play a disproportionate role in disease transmission, which is more consistent with our understanding of sexual networks. Highly connected nodes within scale-free networks also are of fundamental interest to the immunobiology of STIs. The individuals who occupy such nodes are not only highly connected and, therefore, involved in STI transmission, but they also are most likely to acquire immune responses to STI-causing pathogens and to have coinfection with multiple pathogens. Individuals who occupy such nodes in the network are likely to disproportionately shape STI transmission dynamics and pathogen attributes. These individuals can be conceptualized as part of the STI core group.

The third level of network analysis is concerned with how networks themselves change over time. This is particularly relevant to social networks in which structure is not static and is likely to evolve in size and complexity over time, perhaps in response to control programs. Understanding this evolution has remained a major intellectual challenge and has only recently been addressed at a theoretical level in the literature [8, 9, 15]. The most desirable scenario for modeling STI transmission is to incorporate into an integrated template both aspects of an STI’s transmission dynamics—that is, the spread of disease in a static network and the dynamics of the network itself [16, 17].

**IMMUNOBIOLOGICAL ASPECTS OF TRANSMISSION VARIABILITY**

If network analysis is to be used to explore immunobiological influences on STI transmission dynamics, what immunobiological characteristics would be of greatest interest to study? Ultimately, only 2 sources underpin the variability that determines transmission dynamics for STI. These are inherited factors related to the uniqueness of each individual and pathogen genome and the environmental factors related to the precise pattern of encounters with STI that each host experiences and that, in turn, generates adaptive immune responses. Major genetic features relevant to host immunity include different combinations of major histocompatibility complex (MHC) genes, which determine the T lymphocyte repertoire, and polymorphisms in promoters of key regulatory cytokines, the activity of which shapes the polarization and differentiation of the immune response. As important as genomic variability in MHC alleles is to shaping the naive T cell repertoire, a person’s precise history of exposure to STI-causing pathogens is central to shaping the repertoire of memory and effector T and B lymphocytes. The interface between genetics and the behavioral variables that determine exposure to STIs is seen most clearly at the level of immunological memory.

**Effects of host immunogenetics on transmission dynamics.**

The adaptive immune response is strongly influenced by differential activation of T helper (Th) cell subsets. Th1-like cells,
which produce interleukin (IL)–2 and interferon (IFN)–γ, mediate cellular immune responses, whereas Th2-like cells, which produce IL-4, IL-5, IL-10, and IL-13, facilitate humoral immune responses. The pattern of in vivo production of cytokines that follows infection is strongly correlated with disease progression or resolution for several infectious diseases, including STIs [18, 19]. Host genetics are among the several factors that regulate differentiation and polarization to a Th1 or Th2 adaptive immune response and are likely to influence the duration of STI infectivity. Such genetic factors include polymorphisms in the MHC promoter region and genetic differences in the regulation of cytokines [20–22]. For example, secretion of IFN-γ by Th1 cells limits Chlamydia trachomatis infection by depletion of intracellular tryptophan and iron; thus, genetic factors that influence levels of IFN-γ secretion are likely to be important in influencing both the duration of infection and the infectivity of chlamydial disease [23, 24].

IL-10 is another critical cytokine produced by several cell types, including a recently discovered third subtype of T lymphocytes, called “regulatory T cells.” Regulatory T cells have both activating and inhibitory effects on Th1 and Th2 immunity and may promote persistence of STIs. Studies of differential production of IL-10, in relation to immunobiology, are just beginning. For example, Shin et al. [25] recently demonstrated genetic restriction of progression of HIV-1 infection to AIDS, by certain IL-10 promoter alleles. Polymorphisms in IL-10 promoter alleles that result in increased production of IL-10 have been associated with ocular chlamydial infection [26] and also may be important in IL-10–secreting T cell lines that recognize chlamydial antigens and that are commonly identified in fallopian tube tissue from women with tubal factor infertility [27].

Effects of infection history on transmission dynamics. In addition to genetic effects, behavioral variability that results in frequent exposure to STIs also alters transmission dynamics. For instance, an uninfected individual may have partial immunity to a pathogen because of past infection with a related strain or organism. This perhaps is best studied with herpes, because a history of herpes simplex virus (HSV) type 1 infection confers partial cross-protective immunity to HSV-2 infection. The effects of cross-protective immunity to C. trachomatis and Neisseria gonorrhoeae will be more fully explained in the next section.

It is less appreciated that cross-reactive immune responses also can enhance susceptibility to infection, such as when original antigenic sin interferes with development of a specific adaptive immune response, thereby increasing the duration of infection [28]. Original antigenic sin occurs when a second infection with a variant strain of a pathogen stimulates an immune response that is specific to the original variant and that may result in impaired clearance of the second variant. In this situation, immunological memory functions paradoxically to increase the duration of infection and infectivity, by silencing responses to type-specific antigenic sites that engender protective immunity. Original antigenic sin has been well recognized in antibody responses to the C. trachomatis major outer-membrane protein (MOMP) [29].

Some pathogens directly facilitate reinfection through the development of blocking antibodies in response to a primary infection. Blocking antibodies run interference between bacterial antigen epitopes and protective antibodies and T cell receptors, a phenomenon of immunobiological relevance during a repeat infection. In this case, the immune response is directed to a conserved, rather than a variant, antigen. The role of blocking antibodies in repeat gonococcal infection illustrates this important mechanism for pathogen subversion: women with antibodies to gonococcal reduction modifiable protein (Rmp) were more likely to be reinfected than women who lacked Rmp antibodies [30]. Individuals who are frequently exposed and who occupy highly connected nodes in a sexual network are most likely to display immune-memory effects such as cross-protective immunity, original antigenic sin, and blocking-antibody responses.

DO CORE GROUPS PRESENT PATHOGENS WITH UNIQUE IMMUNOLOGICAL ENVIRONMENTS?

The above observations suggest that individuals who are frequently exposed to STIs may have unexpected immunobiological characteristics. As described above, the nonnormal distribution of sexual behavior is compatible with an underlying scale-free network structure of sexual partnerships, in which some individuals have many sexual partnership connections but the majority of individuals have 1 or very few connections. In practice, such highly connected nodes can be conceptualized as the core-group individuals.

An important implication of the core-group concept is that certain STI-causing pathogens must successfully infect highly sexually active individuals to sustain transmission over time. We focus here on whether highly sexually active individuals are likely to have unique immunological characteristics, in comparison with other, less sexually active individuals. Specifically, might there be differences in duration and intensity of infectivity, development and duration of immunity, and susceptibility to infection?

We focus the discussion on data from female commercial sex workers (CSWs) and their clients, because, for CSWs, high rates of partner change have been documented, and numerous epidemiological studies of STI prevalence are available. CSWs are among the most easily studied core-group members, and the immunoepidemiology of STIs among CSWs is likely to
reflect the processes that operate more generally within all core groups.

In several studies, astounding high prevalences of STIs have been reported among CSWs in third-world countries [31–33], and the staggering burden of disease borne by CSWs raises several questions. For example, how does an STI-causing pathogen achieve ecological success in an environment where reinfection and coinfection are the rule? Specifically, how are STIs able to generate productive chains of infection transmission within groups when immune and inflammatory responses are so strongly concentrated?

Immune evasion through antigenic variation: insights from chlamydial and gonococcal infection. For N. gonorrhoeae and C. trachomatis, immune evasion through antigenic variation may be key to persistence in frequently exposed, highly sexually active groups. Antigenic variation extends the duration of infection in a host and facilitates reinfection of a previously exposed host [28, 34]. This strategy is eminently suited to pathogens likely to reencounter hosts.

In a given host-pathogen interaction, an original pathogen strain stimulates a host adaptive immune response against its dominant antigens. Pathogens express variant antigens through genetic processes, such as random mutation generated during DNA replication, genome-encoded antigenic variation, and homologous or heterologous recombination, and escape variants are selected by the immune response. The pace of antigen switching is critical and must be slow enough to allow an individual host to mount a specific response to the dominant antigen but fast enough to prevent elimination of the parasite population as a result of that response. The degree of immunodominance manifested in a given host also determines the strength of selection for antigenic variation in the pathogen [28]. At the population level, escape variants can play an important role in transmission. An inoculum containing multiple antigenic variants has an increased probability of presenting a new variant to an immunologically experienced host.

Thus, antigenic variation contributes to the reproductive number of the pathogen in 2 distinct ways. First, the duration of infection increases, which may be a particularly important determinant of persistence of STIs when rates of partner change are relatively low. Second, antigenic variation facilitates reinfection, particularly in groups of highly sexually active individuals, such as CSWs and their clients, who may already have strain-specific immunity to a pathogen [35, 36].

Host-pathogen interactions characterized by pathogen antigenic variation also tend to generate distinctly different patterns of strain distribution over time. In one case, an antigenic variant may occasionally spread epidemically through the host population. If strain-specific immunity is almost complete and enduring, a large fraction of the host population will be resistant at recovery, and the frequency of the variant will diminish. Several factors determine the pace of emergence of new variants: age structure of the population and the rate at which new susceptible individuals are added to the population, the speed with which variants are spread and cleared, and the waiting time until a potentially successful new variant arises. In such a single-strain system, the most transmissible strain will outcompete other strains [37].

In other cases, multiple antigenic variants may be maintained endemically in a host population. Pathogen populations will organize themselves into strains with nonoverlapping repertoires of antigenic variants if host-population immune responses exert strong selective pressure against the polymorphic antigens. Strain structure also can reflect positive selection for strains that protect each other through the effects of original antigenic sin. Structure also will be influenced by the rate at which new susceptible hosts enter the population, by stochastic processes governing the origin and extinction of strains, and by heterogeneity in host immune responses [38, 39]. Thus, immune selection of pathogen antigens can produce both spatial and temporal variation in transmissibility.

A handful of empirical studies that measure strain structure and patterns of immunity among cohorts of African CSWs are available. These studies do provide supportive evidence regarding a discrete strain structure in C. trachomatis and N. gonorrhoeae populations [35, 36]. In a 2.5-year study of CSWs in Nairobi, 18 different chlamydial genotypes were observed. Significant annual shifts in MOMP allelic composition in the pathogen population were observed over the 2.5-year period, with several variants disappearing from and several new variants arising in the pathogen population. These dynamic fluctuations in the strain-specific composition of the C. trachomatis population and the observation that the majority of reinfections, spaced at least 6 months apart, were due to infection with different C. trachomatis MOMP genotypes were consistent with a process of strain-specific immunity structuring pathogen populations into antigenically discrete strains [36].

In addition, discrete strain structure and evidence of strain-specific immunity to N. gonorrhoeae also has been reported in a 1-year study of CSWs in Nairobi [35]. Gonococcal isolates were typed according to protein 1 serovars. Four strains accounted for 70% of all serovars, during the initial 7 months of the study. In the second half of the study, a few strains continued to account for the majority of infections; however, their frequencies had changed dramatically.

The observed pattern in both studies was consistent with cyclical or nonperiodic fluctuation between sets of strains produced by heterogeneity in host immune response, although immune responses were not directly measured. These studies are unusual in uncovering the epidemiological pattern of strain variation and substitution, in part because they involved CSWs. Such patterns for chlamydial and gonococcal strain variation
are much less readily discerned when studies are conducted in the general population, in which immune-selection forces are expected to be much less intense.

Transmission dynamics and potential effects on evolution toward and away from virulence. The evolution of STI-causing pathogens is determined, in part, by the effect that immunity exerts on pathogen virulence and by the relationship between virulence and transmission success. The immune response provides an ultimate limit on transmission, through clearance of infection; yet, the process through which this occurs may involve the induction of an inflammatory response that actually enhances transmission via the development of an infectious exudate or lesion. In the simple case in which virulence is inversely related to duration of infection, STI-causing pathogens that circulate within highly sexually active groups tend to evolve toward virulence, and those that circulate within less sexually active groups tend to evolve toward avirulence. For instance, gonococcal strains with nutritional auxotrophy, such as AHU⁻ (arginine, hypoxanthine, uracil-requiring) or PCU⁻ (proline, citrulline, uracil-requiring), induce little inflammation and appear to cause long-term infection (6 months), which may be an adaptation for circulation in less sexually active groups. Gonococcal strains with no nutritional requirements are prototrophs. They induce greater inflammation and are commonly recovered from highly sexually active groups. Gonococcal strains with no nutritional requirements are prototrophs. They induce greater inflammation and are commonly recovered from highly sexually active groups. Similar adaptations also may be apparent for chlamydial strains: virulent L serovars have been found mainly among CSWs, and non-L serovars have been found to be distributed in the general population [36].

Transmission dynamics and pathogen gene exchange. We have drawn attention to the presence of dynamic strain structure in populations of *N. gonorrhoeae* and *C. trachomatis*. Genetic processes maintaining strain structures in these organisms are very different. Antigen switches in gonococcal strains can be achieved by homologous recombination, since gonococcal strains are transformation competent. Thus, horizontal gene transfer enables individual gonococcal cells to acquire beneficial genetic material from neighboring cells before internal switching and clonal expansion have occurred [40]. In contrast, chlamydial strains are incompetent for facile horizontal DNA exchange, and only limited evidence has been found for horizontal gene transfer into chlamydial genomes. Antigenic variation appears to be limited to the variable regions of the MOMP, and genomic analyses indicate limited intrinsic mechanisms for recombination. Yet, horizontal gene transfer seems to be occurring, since analysis of chlamydial isolates from a cohort of CSWs in Nairobi provided evidence for substantial MOMP mosaicism that suggests recombination [41]. Coinfection of core-group members may be the unique setting in which gene exchange can occur and may have broader relevance to gene exchange among other STI-causing pathogens. Thus, new genetic variants of an STI-causing pathogen may have their origin in core-group members in part through coinfection and immune-selection processes.

In aggregate, these observations, although limited, do suggest that core-group members present STI-causing pathogens with unique immunological environments. Such environments appear to trigger the evolution of immune-evasion mechanisms, antigenic diversification, and virulence properties that typify sexually transmitted microbes.

CONCLUSION

We have discussed 2 important features that partially determine how STIs persist and evolve: the unique immune responses related to each host and the environmental factors related to the precise pattern of pathogen encounters that each host experiences. Both of these aspects are independently the subject of intense study spanning disciplines as diverse as immunology and sociology. However, the study of such complex systems requires theoretical tools that incorporate all such features into an integrated template, to estimate the impact of each on the global pattern of STI transmission within a population. When developed, these templates should prove useful for the prediction of future trends in STI epidemiology.

Mathematical models are powerful means to generate new theoretical tools. Inconsistency between certain theoretical results and real data suggests that more conceptual elements need to be brought into the mathematical models. Of importance is the need to incorporate heterogeneity in the transmissibility of infection. Thus far, this parameter has been taken either as an average value for a population, which neglects individual characteristics and, therefore, the variability in transmission between individuals, or from a fixed probability distribution, which neglects the variability in susceptibility and infectiousness within an individual over time. Thus, models need to address in explicit detail the transmission variability related to the global structure of sexual networks. Early study has suggested that network topology offers the critical framework in which to evaluate the conditions for STI evolution and persistence. In future, mapping the topology of sexual networks will be important, including consideration of the fact that the networks themselves are likely to evolve over time.

Advances in genomics now provide the tools to conduct much-needed immunoepidemiological studies. The challenge is to understand how molecular and cellular networks that operate within a single host, in response to STIs, are embedded within and interact with the social networks that transmit STIs at the population level. Remarkable recent work by Barabasi et al. [42] has shown that network structures that operate at the molecular level, through genome expression, and metabolic networks have the same scale-free structure as networks that operate at the social and ecological levels. The immunobiological networks that operate in STIs within individuals may
have a topology similar to that of the social networks that transmit STIs at the population level. Ultimately, such studies should yield the insights and tools necessary to move public health efforts from the partial control of STIs to more-complete prevention and elimination.

References