STD Transmission Dynamics: Some Current Complexities

2002 Thomas Parran Award Lecture

RICHARD ROTHENBERG, MD, MPH

MY DEEPEST THANKS to the Association and to the Award Committee for my having been chosen to receive the Parran Award. My special thanks and appreciation to John Potterat, a colleague and collaborator for 30 years, and no less fervent thanks to my family, friends, and associates—the many people who have created the self-organizing microcosm that I have been privileged to inhabit. Receiving a "lifetime" award is daunting for two reasons: first, to contemplate that it has been a lifetime; second, to realize that the lifetime has focused on a field in its infancy, a field that keeps asking better and better questions. . .

Pursuing the dynamics of disease transmission requires the transformation of a fundamental epidemiologic question. In epidemiology, we wish to know the comparative effect of a risk: How disease occurrence in persons with a risk compares with occurrence in persons without the risk. More generally, we look for etiologic insight into disease processes by measuring the adjusted (unconfounded) relative risks for factors associated with those processes. A complex, but increasingly coherent, set of logical principles has grown up around this question.[1] For transmission dynamics, we pose different questions: Given a set of risks (or a set of starting conditions), what will happen? How much disease will be transmitted, by whom, to whom, and with what sort of trajectory? What makes transmission start, or stop, or stay constant? How do the traditional measures of association (e.g., relative risk, attributable risk, transmission probability) help us understand what happens in populations?

Relative Risk

A number of investigators have examined the route from risks to incidence. In his early consideration of the use of the odds ratio (OR) to estimate relative risks from case-control data in diseases of low prevalence (the "rare disease" assumption), Cornfield[2] demonstrated the simple algebraic relationship between the probability of exposure, the probability (prevalence) of disease, and the occurrence of disease per unit population. Other investigators have explored methods for combining crude incidence in demographic strata with relative risks to estimate exposure specific incidence,[3–6] an extension of techniques devised to estimate the probability of developing a disease.[7] Lele and Whittemore[8] have explored the question of how much of the excess disease in a population can be attributed to risk factor configuration. These approaches have been developed in the context of chronic diseases, such as cancer, and depend on prior knowledge of crude incidence. Predicting incidence from relative risks, especially in the context of infectious diseases that require knowledge of contact patterns and transmission dynamics, is more difficult.

Consider, for example, the relationship of a known set of risk factors to disease incidence. In principle, there is a straightforward connection. Incidence in a population (I_T) may be thought of as a weighted average of incidence in persons exposed to a risk (I_E) and incidence in persons who are unexposed (I_U), with the proportion of each in the population acting as the weights: $I_T = I_E \cdot p(E) + I_U \cdot p(U)$. Since the $RR = I_E/I_U$, then $I_T = (I_U)[(p_U) + RR(p_E)]$, and, in the case of multiple risks, $I_T = \sum_i \{(I_{Ui})[(p_{Ui}) + RR_i(p_{Ei})]\}$. The latter formulation, however, cannot be applied directly, because it ignores the fact that it creates multiple sectors in the population—groups with different combinations of risks. Each mutually exclusive piece of the population is subjected to a specific combination of risks whose interaction provides the overall relative risk for that subset. Such interaction may be additive, synergistic, or antagonistic, in ways that may not be predictable a priori. The comparison group for that subset (and every subset) is the group of persons who are unexposed to all risks (not the group that is unexposed to a specific combination). Theoretically, summation over these pieces would predict incidence in the population, but a number of empirical factors mitigate against such an approach:[1] it would not account for multiple levels of risk nor the dynamics of change in risk;[2] it does not account for group or population level factors (network characteristics, for example) that can have independent influence on transmission;[3] perhaps most important, it does not take into account the biology of the organism or host-organism interactions.

Thus, the theoretical potential of deriving incidence from relative risks is compromised by the empirical complexity of assessing and combining risks. There may be theoretical concerns as well. The expression of each of the putative risks is associated with measured and unmeasured variability that can have unpredictable effects on the outcome of interest. The use of summary measures of association may provide an estimate of incidence, but does not furnish insight into the intermediate dynamics. Finally, an interesting group, usually ignored in epidemiologic studies, is the lynchpin for prediction. The incidence in the unexposed (I_U) is the scaling factor that determines the final level of incidence, and the relative risk approach provides no insight,

From the Department of Family and Preventive Medicine, Emory University School of Medicine, Atlanta, Georgia

Correspondence: Richard Rothenberg, MD, MPH, Emory University School of Medicine, Department of Family and Preventive Medicine, 69 Butler St., SE, Atlanta, GA 30303. Email: rrothen@emory.edu

Received for publication November 27, 2002 and accepted December 30, 2002.

TABLE 1. Multiplex Distribution of Contacts, Atlanta, GA, 1995–1999

Type of Contact	No. of Contacts: Sexual	Social	Drug	Needle	Combinations
Sexual	171	708	470	5	
Social		1712	638	17	
Drug			344	18	
Needle				0	
Sexual, social, drug					254
Sexual, social, needle					5
Sexual, drug, needle					4
Social, drug, needle					15
Sexual, social, drug, needle					4
Total contacts					4365

theoretical or empirical, into the incidence in persons for whom no risk factors are identified.

Attributable Risk

A similar reasoning applies to the measure that associates a particular risk factor with a definable proportion of disease occurrence. The classic formulation, often referred to as Population Attributable Risk (PAR) derived by Levin[9] in 1953, is

$$\mathrm{PAR} = \frac{p_e(\mathrm{RR} - 1)}{1 + p_e(\mathrm{RR} - 1)}.$$

Use of this univariate approach, in the presence of multiple risks, can easily result in a PAR more than 100%. To deal with multiple risks, in a manner analogous to combining relative risks, the formulation is

$$\mathrm{PAR} = \frac{\sum p_{ei}(\mathrm{aRR_i} - 1)}{1 + \sum p_{ei}(\mathrm{aRR_i} - 1)},$$

where *a* denotes adjustment of the RR for confounding.[10] To calculate the overall PAR, all possible combinations of risks must be defined (for example, for n risks, there are 2^n combinations including the referent [no risk] group) and both the relative risk (RR_i) and population exposure proportion (p_{ei}) determined for each group. From these, the $\sum p_{ei}(\mathrm{aRR_i}-1)$ can be determined. Here, an important theoretical issue arises. Since each category involves a combination of one or more risks that cannot be further subdivided, summing over a single risk to determine its attributable portion is not possible. Thus, this hallmark of attribution in epidemiology is not really available except for situations in which only one risk is important. This theoretical constraint to understanding how diseases are transmitted is compounded by the empirical difficulties of collecting comprehensive data and determining adjusted estimates for each combination of risks.

Transmission Probabilities

A number of investigators[11–13] have attempted to examine dynamics by using a Bernoulli approach to calculate the probability of transmission. Generically, the probability of transmission of infection from *k* to *i* would be represented as

$$P(\mathrm{inf})_{ki} = [1 - \Pi_j (1 - f_j)^{c_{ijk}}] \, Pr_k,$$

where *f* is the estimated probability of each type of act, *c* is the number of acts of type *j* between *k* and *i*, and Pr_k is the prevalence of infection in the population from which *k* comes. This approach poses empirical challenges because of the need for data on types and frequencies of each kind of risky act, knowledge of time intervals, and adequate estimates of transmission probabilities. Theoretically, because *I* is probably not *K's* only partner within some time interval, the inseparability of multiple risks delivered through multiple channels renders attribution of disease transmission to a single type of act, or to an individual (the "source," in STD parlance), moot. This indivisibility of a constellation of risks reduces the investigator's ability to define the dynamics of transmission in the traditional terms of univariate, hierarchical risk.

A number of different data approaches support the contention that multiple risks coexist, particularly in populations deemed at high risk for STD or HIV transmission. For example, in a study of inner city men and women at risk for HIV, there is a striking multiplexity of risks taken with contacts identified by the study's primary respondents (Table 1).[14] Of the 4365 contacts reported by participants in this study, 51% involved only a single type of relationship, but only 12% involved sex or drug use alone (no dyad reported needle-sharing as the only risk). Combinations of risk were far more likely. The same message is conveyed from a different population in the same setting (Table 2).[15] Persons with HIV, recruited from the community and not in treatment, were initially classified by traditional categories (MSM, male IDU, Female). It is apparent that this classification ignores the considerable multiplicity of risks within each group. For example, 22% of men classified as IDU had sex with other men, men as well as women exchanged sex for drugs or money with substantial frequency, and none of the groups used condoms regularly for any type of sexual activity.

The multiplicity of risk is complicated by the preexisting prevalences of multiple diseases in some populations. In an ongoing study of STD transmission among adolescents (R.R., unpublished data, 2003), the frequencies of STDs among the contacts (social, sexual, or drug-using) to persons who do *not* have a particular STD are high because they reflect the background prevalences in that population (Table 3). The notion of screening the contacts of negative persons and finding a high proportion with STD is certainly a novel one for STD control and not likely to take root in the near future. The phenomenon is striking, however, because it points to the complex interrelationships among multiple diseases that circulate simultaneously, but differently, in populations. Of note is the additional observation that none of these adolescents had HIV or syphilis, and only two persons (older contacts to two different respondents) had antibody to HCV.

TABLE 2. Frequency of Risk Taking Among Persons With HIV Recruited From the Community and not in Formal Treatment Programs, Atlanta, GA, 1998–2001

	Community Persons, HIV+ (%)		
	MSM	IDU	Female
Same sex sexual orientation	100	22	11
Sex with men	90	28	89
Sex with women	14	72	5
Sex with a crack user	57	83	79
Sex with an IDU	19	11	21
Sex with a heroin user	10	17	5
Given money for sex	10	17	5
Given drugs for sex	29	56	0
Received money for sex	33	17	47
Received drugs for sex	33	28	42
Receptive anal intercourse	76	17	21
Active anal intercourse	67	22	0
Condom use with anal sex	0	0	0
Condom use with oral sex	0	0	0
Condom use with vag sex	0	0	0
Condom use at last sex	38	39	37

Newer Approaches

Networks

Although not new, network theory and analysis have only recently been applied to the problems of infectious disease transmission.[16] Intense empirical and theoretical development have been rewarding, but have also uncovered the fault lines in both approaches. As with all empirical studies, missing data pose a problem, but this problem may be of particular significance in network studies, since the failure to identify critical links between persons can distort both analytical and visual results. Similarly, sampling methods can modify the observed outcome, and it is not yet clear what influence specific study designs have on observed network configuration. The fact that many of the populations of interest for disease transmission are "hidden," or "elusive" precludes traditional random sampling designs,[17,18] with an uncertain effect on estimation and generalizability.

Theoretical development requires some simplification of the complex, and often unknown, rules that govern processes. Some assumptions used in theoretical development and simulation (such as homogeneity, or random contact) can provide interesting results, but may not offer a conduit to real situations. In addition, the estimates used in simulations are usually the result of theory-based estimates themselves, rather than of measurement, and provide for a degree of uncertainty that is often ignored. Part of the problem may be the heightened expectations imposed on theoreticians and disease modelers. Such a burden undercuts the fundamental purpose of the exercise: not to produce "the truth," but to examine the complexity in manageable terms.

The fault lines in each approach emphasize the increasing need for their rapprochement. One such effort[19] uses a large set of completed empirical studies to inform the assumptions and estimates for simulations generated by the Markov Chain Monte Carlo (MCMC) approach (a well-established statistical method more recently adapted for data generation) and analyzed using "p*" methods (logistic regression that models network outcomes as a function of network structure). The result of this effort, still in its early stages, will be to provide a method for combining empirical and theoretical results, and hopefully to provide insight into the effect of network structure on disease propagation.

Newer Theory

Another auspicious merger of empirical data with network theory has emerged from contemplation of the World Wide Web, cellular genetics, and other large networks. Statistical physicists and others have reexamined graph theory in light of the huge data sources that describe social organization (such as the Web) as well as a variety of other large-scale networks that occur in biologic systems. Two interesting properties of these types of networks have emerged: scale-free distributions and small-world phenomena.[20] The scale-free properties—a termed coined as an "analogy with fractals, phase transitions, and other situations where power laws arise and no single characteristic scale can be defined"[20]—of these large webs stem from the nature of the degree distribution of nodes. Over a wide range, the probability of observing a given degree (P(x), where x is the number of contacts to a node follows a power law distribution, $P(x) \sim x^{-\gamma}$, and the exponent, γ, is usually between 2 and 3. This configuration has been observed for the Web itself[21–23] as well as congeners such as instant messaging;[24] for scientific collaboration, for movie actor collaboration, for service on the boards of Fortune 500 companies;[25] for the spatial distribution of earthquakes;[26] for the adoption of faith-based systems in societies;[27] and for electrical power grids, although for several of these, a somewhat better fit is possible with

TABLE 3. Prevalence of STD Infections in Contacts to Adolescents Who Tested Negative for These STDs, Atlanta, GA, 2000–2002

			Proportion of Contacts Who Were Positive for Each of These Infections						
		N	HBV	HCV	HSV1	HSV2	GC	CT	TR
Index cases who tested negative for each of these infections	HBV	102	3.9	1.0	39.2	35.3	5.9	15.7	2.9
	HCV	119	9.2	0.8	38.7	34.5	5.9	14.3	5.0
	HSV1	48	6.3	0.0	10.4	10.4	4.2	8.3	0.0
	HSV2	61	9.8	1.6	21.3	16.4	3.3	8.2	1.6
	GC	110	10.0	0.0	39.1	33.6	2.7	12.7	5.5
	CT	99	10.1	0.0	35.4	32.3	4.0	10.1	6.1
	TR	116	7.8	0.9	38.8	34.5	6.0	16.4	3.4
Prevalence of infection in overall population*		300	14.0	0.7	53.0	40.7	7.0	17.7	4.7

*Represents the prevalence in the entire cohort screened to date.

alternate models.[20] Most interesting from the point of view of sexual transmission is that, in at least one large survey of sexual activity, the numbers of sexual partners also appears to follow a scale-free distribution.[28]

Detailed study of the scale-free distribution has been facilitated by the development of a model, incorporating both growth and preferential attachment, that simulates the structure of the Web.[21] Based on such a model, scale-free networks exhibit significant clustering, are resistant to random "attack" (that is, dissolution from random removal of nodes) but may well be susceptible to targeted attack on high degree nodes.[29] The nature of this clustering, the result of assortative mixing by degree, has been explored in detail and new measures of assortativeness introduced.[30,31] Scale-free networks do not appear to be subject to an epidemic threshold for continuing spread of disease,[32,33] with certain fiats related to the analytic approach.[34] In comparison, networks simulated using a "deactivation model," generate a collection of "stars" (that is, central nodes with their set of contacts) that are connected in a chain, and do exhibit an epidemic threshold.[35,36]

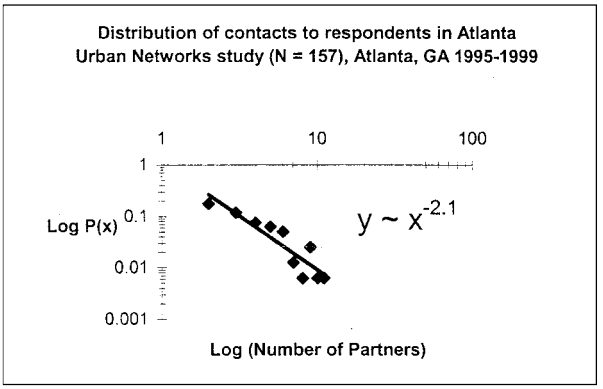
The connection of these findings to empirical networks that experience STD/HIV transmission are still being explored, but it is interesting to point out that the scale-free networks correspond to those that have been termed "cyclic," in network analysis, and the deactivation-generated networks resemble those that have been called "dendritic."[37–39] Because the former may not have an epidemic threshold and the latter do is congruent with current notions of endemicity in such populations. Cyclic networks are likely to support transmission after an introduction; dendritic networks are more likely to require larger, or repeated, introductions of a disease agent to establish endemic transmission.

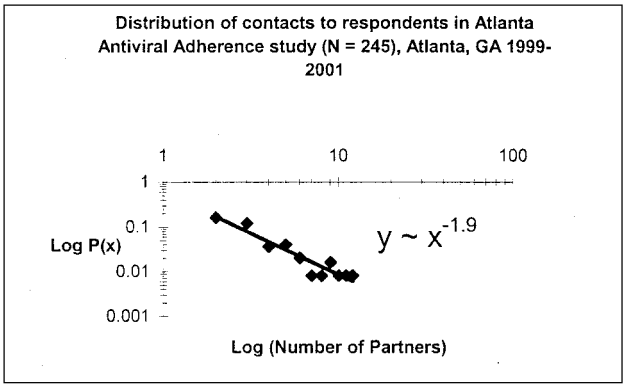
In addition, scale-free properties associated with giant networks may be observed—albeit with less stable estimates—in the small groups that are studied empirically. Each of the three studies in Atlanta cited earlier (see Tables 1 to 3) contains groups of respondents who identify varying numbers of sexual partners. Each of these has degree distributions that can be described by a power law with exponents of approximately 2.0 (Figure 1). It should be noted, however, that in some cases, the power law curve may not be the best "fit" to the data,[20,40] and some alternative mechanisms can produce the markedly skewed distribution that characterizes the scale-free phenomena. This observation suggests that the self-organizing principles that appear to govern many large networks may apply to smaller, local networks as well. If the "unseen hand" of network formation operates without regard to network size (at least approximately), important insights into disease transmission dynamics may be available from a wide range of biologic, mechanical, and social systems.

Complexity

One of those insights may be a better understanding of sexual transmission as a complex system.[41] STDs and HIV, in their population epidemiology, have many of the hallmarks of a complex system:

Sensitivity to starting conditions. It is apparent from small-scale empirical studies that there is a correlation between structure and the extent to which disease propagates. Structure results from personal behavioral choice, and from the type and frequency of sexual encounters within a network. The latter are subject to considerable (perhaps unmeasurable) variability, so that the resulting structure may be subject to small changes that have considerable impact on transmission. Modeling efforts, such as those described earlier may need to take such unmeasured variability into account in deriving a more quantitative relationship between structure and transmission.





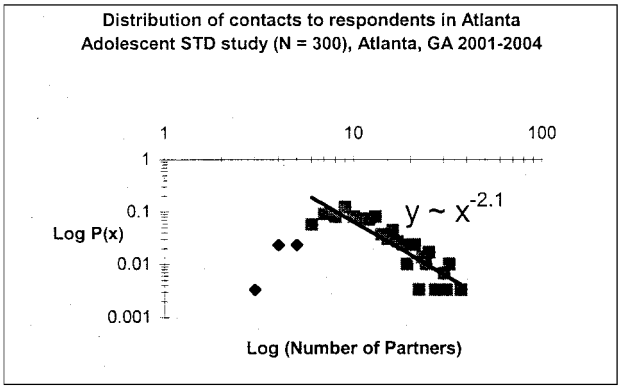


Fig. 1. Degree distribution of sexual partners in three studies of STD/HIV transmission in Atlanta, Georgia 1995 to 2001.

Scale invariant (fractal) behavior. Strict geometric fractals (the type associated with striking pictures of patterns whose large and small scale geometry is the same) are probably less common than statistical fractals, wherein the reproduction of patterns at different levels of measurement has similar statistical properties. The overall pattern of STD distribution is typified, nationally, by focal aggregation of cases in selected states. Each state, in turn, exhibits focal aggregation in its cities; the cities, in certain census tracts (or other jurisdictional markers); the census tracts, in neighborhoods.[42] This pattern of "hot spots" reflects a similarly heterogeneous distribution (long tail to the right) of cases, whatever the metric. Such a distribution has a metaphoric correspondence with

degree distribution, through the existence of "nodes" of heightened activity that serve as foci for continued disease propagation.

Bifurcations and nonrepeating periodicities. The occurrence of STDs has long been known to follow a sawtooth pattern whose fluctuations are linked to seasonal change. Such (relatively) sudden and distinct oscillations are also characteristic of chaotic systems that experience phase transitions. Superimposed on these short-term oscillations are longer-term trends that are resistant to a priori prediction, but for which facile post hoc explanations are usually available. Such aleatoric periodicities are, however, characteristic of chaotic systems.

Self-organization of complex patterns. As noted, empirical studies have documented the approximate correlation of network configuration with disease endemicity. Although considerable work is needed to place these observations on firmer empirical and theoretical footing, the current indications suggest that the personal choices of individuals self-organize to create network structures that are associated with specific levels of disease transmission. If the parallels with large social and biologic networks persist, such self-organization may be an important marker of complex processes in STD transmission.

These separate streams of thought and inquiry—statistical methods to meld theory and simulation with empirical data; analysis of large network systems; complexity theory—seem to be converging in the arena of STD/HIV transmission, and highlight the vital, roiling nature of the field. This strange attraction between sophisticated mathematical approaches and acts of sex has indeed led to better questions.

References

1. Rothman KJ, Greenland S. Modern Epidemiology. Philadelphia: Lippincott Williams & Wilkins, 1998.
2. Cornfield J. A method of estimating comparative rates from clinical data. Applications to cancer of the lung, breast, and cervix. JNCI 1951; 11:1269–75.
3. Benichou J, Wacholder S. A comparison of three approaches to estimate exposure-specific incidence rates from population-base case-control data. Stat Med 1994; 13:651–61.
4. Benichou J, Gail MH. Methods of inference for estimates of absolute risk derived from population-based case-control studies. Biometrics 1995; 51:182–94.
5. Greenland S. Multivariate estimation of exposure-specific incidence from case-control studies. J Chronic Dis 1981; 34:445–53.
6. Neutra RR, Drolette ME. Estimating exposure-specific disease rates from case-control studies using Bayes' theorem. Am J Epidemiol 1978; 99:325–32.
7. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. JNCI 1989; 81:1879–86.
8. Lele C, Whittemore AS. Different disease rates in two populations: how much is due to differences in risk factors? Stat Med 1997; 16:2543–54.
9. Levin ML. The occurrence of lung cance in man. Acta Union Intern Cancer 1953; 9:531–41.
10. Walter SD. The estimation and interpretation of attributable risk in health research. Biometrics 1976; 32:829–49.
11. Allard R. A mathematical model to describe the risk of infection from sharing injection equipment. J Acquir Immun Defic Syndr 1990; 3:1010–1016.
12. Bell DC, Trevino RA. Modeling HIV risk. J Acquir Immun Defic Syndr 1999; 22:280–287.
13. Kaplan EH. Modeling HIV infectivity: must sex acts be counted? J Acquir Immun Defic Syndr 1990; 3:55–61.
14. Rothenberg RB, Long D, Sterk C, et al. The Atlanta urban networks study: a blueprint for endemic transmission. AIDS 2000; 14:2191–200.
15. Rothenberg RB, Baldwin JA, Trotter R, Muth SQ. The risk environment for HIV transmission: results from the Atlanta and Flagstaff network studies. J Urban Health 2001; 78:419–32.
16. Klovdahl AS. Social networks and the spread of infectious diseases: the AIDS example. Soc Sci Med 1985; 21:1203–16.
17. Rothenberg RB. Commentary. Sampling in social networks. Connections 1995; 18:105–11.
18. Watters JK, Biernacki P. Targeted sampling: options for the study of hidden populations. Soc Probl 1989; 36:416–30.
19. Morris, M. Local Rules and global properties: modeling the emergence of network structure. Workshop on dynamic social network analysis. 2002. National Academy of Sciences, National Research Council Committee on Human Factors.
20. Strogatz SH. Exploring complex networks. Nature 2001; 410:268–76.
21. Barabasi AL, Albert R. Emergence of scaling in random networks. Science 1999; 286:509–12.
22. Barabasi AL, Albert R, Jeong H, Bianconi.G. Power law distribution of the World Wide Web. Science 2000; 287:2115a.
23. Barabasi AL, Albert R, Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web. Physica A 2000; 281:69–77.
24. Smith, R. Instant messaging as a scale-free network. arXiv:cond-mat/0210146. June 19, 2002. http://www.arXiv.org.
25. Newman MEJ, Watts DJ, Strogatz SH. Random graph models of social networks. PNAS 2002; 99:2566–72.
26. Abe, S, Suzuki, N. Scale-free network of earthquakes. arXiv:cond-mat/0210289. October 13, 2002. http://www.arXiv.org.
27. Sadedin, S, Dybiec, B, Briscoe, G. A toy model of fairth-based systems evolution. arXiv:cond-mat/0210292. October 14, 2002. http://arXiv.com.
28. Liljeros F, Edling CR, Amaral LAN, Stanley HE, Aberg Y. The web of human sexual contacts. Nature 2001; 411:907–8.
29. Vazquez, A, Moreno, Y. Resilience to damage of graphs with degree correlations. arXiv:cond-mat/0210146. September 7, 2002. http://www.arXiv.org.
30. Newman, MEJ, Girvan, M. Mixing patterns and community structure in networks. arXiv:cond-mat/0210146. October 7, 2002. http://arXiv.com.
31. Newman, MEJ. Assortative mixing in networks. arXiv:cond-mat/0205405 v1. May 20, 2002. http://www.arXiv.org.
32. Boguna, M, Pastor-Satorras, R, Vespignani, A. Absence of epidemic threshold in scale-free networks with connectivity correlations. arXiv:cond-mat/0208163 v1. October 14, 2002. http://www.arXiv.org.
33. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Phys Rev Lett 2002; 89:3200–3203.
34. Newman, MEJ. The spread of epidemics on networks. arXiv:cond-mat/0205405 v1. April 30, 2002. http://www.arXiv.org.
35. Eguiluz, VM, Klemm, K. Epidemic threshold in structured scale-free networks. arXiv:cond-mat/0205439 v1. May 21, 2002. http://www.arXiv.org.
36. Moreno, Y, Vazquez, A. Disease spreading in structured scale-free networks. arXiv:cond-mat/0210362 v1. October 17, 2002. http://www.arXiv.org.
37. Jolly AM, Wylie JL. Sampling individuals with large sexual networks: an evaluation of four approaches. Sex Transm Dis 2001; 28:200–207.
38. Potterat JJ, Phillips-Plummer L, Muth SQ, et al. Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs. Sex Transm Infect 2002; 78:i159-i163.
39. Potterat JJ, Muth SQ, Rothenberg RB, et al. Sexual network structure as indicator of epidemic phase. Sex Transm Infect 2002; 78:i152-i158.
40. Jones, JH, Handcock, MS. An assessment of preferential attachment as a mechanism for human sexual network formation. 10–15–2002. Center for Statistics and Social Sciences, University of Washington, Working Paper 23.
41. Newman MEJ. Mixing patterns in networks. Phys Rev E 2003;67: 02612.6.
42. Rothenberg, RB, Potterat, JJ. Gonorrhea Surveillance: the missing links. Sex Transm Dis 2002; 29:806.