Determinants of STD epidemics: implications for phase appropriate intervention strategies

S O Aral

Determinants of evolving epidemics of sexually transmitted diseases (STD) are equally influenced by the evolution of the STD epidemics themselves and by the evolution of human societies. A temporal approach to STD transmission dynamics suggests the need to monitor infectivity, rate of exposure between infected and susceptible individuals, and duration of infectiousness in societies. Different indicators may be used to monitor rate of exposure in the general population and in core groups. In addition, underlying determinants of STD epidemics such as poverty, inequality, racial/ethnic discrimination, unemployment, sex ratio, volume of migration, and health care coverage and quality are important variables to monitor through a surveillance system focused on social context. Ongoing large scale societal changes including urbanisation, globalisation, increasing inequality, and increasing volume of migrant populations may affect the evolution of STD epidemics. Globalised STD epidemics could pose a major challenge to local public health systems.

"Our conception of the determinants of an epidemic changes as the epidemic evolves: at certain times pressure arises to track infection and improve access to health care systems; at other times, the emphasis has been on individual risk factors, risk taking or on "risk groups;" eventually, rising poverty and social deprivation may figure more predominantly in policy making." Atlani L, et al. Social change and HIV in the former USSR: the making of a new epidemic. Social Science and Medicine 2000;50:1547–56.

The complex interactions between the evolution of STD epidemics and our comprehension of such evolution are difficult to grasp. Our understanding is shaped to a great extent by public health pressures to control the epidemic, and the political pressures that originate in cultural sensitivities, power dynamics, and economics of resource allocation. All of these pressures are highly interactive and dynamic. Identification and implementation of effective prevention programmes depend on a scientifically sound, objective understanding of how STD epidemics evolve.

Correspondence to:
Dr Sevgi O Aral, Division of STD Prevention, CDC, 1600 Clifton Rd NE, M/SE02, Atlanta, GA 30333, USA; address for reprints: CDC, NCHSTP, Information Technology Services, 1600 Clifton Rd, M/S E06, Atlanta, GA 30333, USA

Accepted 29 November 2001
population dynamics of infectious diseases and applied to the spread of sexually transmitted infections proposed that the rate of spread of sexually transmitted infections (R0) is determined by the infectivity of the pathogen or the probability of infection in case of exposure (β), the rate of exposure between infected and susceptible individuals (C), and the duration of infectiousness (D). On the basis of this theoretical model, a conceptual framework which differentiates between underlying distant and proximate determinants of STD epidemics has been proposed (fig 2).

Another conceptual framework focused on the dynamic topology of STD epidemics proposed that the evolution of these epidemics through predictable phases is shaped by the dynamic interplay among the pathogen, the behaviours of subpopulations, and the prevention efforts developed to limit the impact of the epidemic. In this approach it is important to identify the determinants of infectivity of the pathogen, behaviours of the population, and prevention efforts. Both the “proximate and underlying determinants of STD epidemics” framework and the “dynamic topology” framework have the advantage of incorporating infectious spread of sexually transmitted pathogens and the evolution of STD epidemics.

On a different note, the recent decades have been remarkable with respect to the evolution of human societies. This era has been marked by major developments in communication technologies and volume of transportation, increasing inequality within societies as well as increasing inequality between countries; by the emergence and growth of globalisation; by an increasing proportion of “edge walkers” (or people who straddle cultures owing to immigration, social mobility, displacement, and so on) globally; by increased proportions of population living in post-conflict societies; by increasing urbanisation; and by decreasing demand for low skilled labour which results in growing proportions of unemployed in otherwise well to do societies. Moreover, in industrialised countries, the institutions of marriage and family have been evolving, with greater proportions of individuals being single as a result of decreasing marriage rates, increasing divorce rates, delayed marriages, increasing life expectancy, and changing gender roles. All these changes in the structure and functioning of

Figure 1  Levels of causation of sexually transmitted diseases (STD) and corresponding types of health intervention.

Figure 2  Determinants of sexually transmitted disease (STD) epidemics: a crude framework.
human societies affect STD epidemics, their evolution, their determinants, and the evolution of these determinants greatly. In this article I will attempt to describe the determinants of STD epidemics and their evolution. There are three major parts to the paper. First, I will discuss issues of operationalisation, measurement, and monitoring of the immediate or most proximate determinants of STD epidemics. These are the three components of the \( R_0 \) model: infectivity (\( \beta \)), rate of contact between infected and susceptible individuals (\( C \)), and duration of infection (\( D \)). Second, I describe some of the most important and large scale global social changes and their effects on STD epidemics. Finally I will consider how the evolution of human societies, as evidenced by the large scale global social changes, may affect the evolution of STD epidemics through predictable epidemic phases.

**IMMEDIATE DETERMINANTS OF STD EPIDEMICS: \( \beta \), \( C \), and \( D \)**

The immediate determinants of STD epidemics include the three components that determine the rate of spread (\( R_0 \)) of sexually transmitted infections in populations: infectivity (\( \beta \)), rate of exposure between infected and susceptible individuals (\( C \)), and duration of infection (\( D \)), and the factors that affect those variables. These determinants are both necessary and sufficient to define the rate of spread of any STI in all populations at all times. The determinants themselves are subject to great heterogeneity across populations, over time, and of course, across sexually transmitted pathogens. This variability requires that we identify parameters that reflect the values of \( \beta \), \( C \), and \( D \) and try to measure them.

**Infectivity**

Research has focused on several variables that influence the value of \( \beta \). These have ranged from use of condoms to viral load in the case of HIV infection, and include the role of other STDs as co-factors, circumcision status, effects of suppressive treatments, and sexual practices such as dry sex. Arguments based on biological plausibility make a strong case for the effect of these factors on the value of \( \beta \); however, accurate measurement of most such effects is yet to be realised.

In a structured review of studies on condom effectiveness in reducing heterosexual HIV transmission, the proportionate reduction in HIV infection with condom use was approximately 80%. Published scientific evidence for condom effectiveness in preventing STDs concluded that condoms prevent transmission of HIV and gonorrhoea, although there was inadequate evidence to draw conclusions about condom effectiveness in the prevention of STDs other than HIV and gonorrhoea. Recent studies have highlighted a number of issues that may affect the evaluation of condoms for STD prevention. These include the potential problems associated with self reported condom use, failure to ask whether condoms were used correctly, and whether slippage or breakage occurred with their use.

Moreover, the organism to organism variation in infectivity appears to be great. This leads to a rather large “forgiveness factor” in case of condom effectiveness for less efficiently transmitted infections like HIV and for those sexually transmitted infections (STIs) that are less prevalent in the population (Cates W, personal communication). These variations in infectivity across sexually transmitted pathogens mean that it is more difficult to demonstrate the effectiveness of condoms in reducing the transmission of more efficiently transmitted bacterial STD. Uptake of condom use in response to public health interventions may pose an additional problem. In a recent study among HSV-2 discordant couples, even with regular counselling only about 10% overall reported consistent (100%) condom use and only about 25% reported using them half the time.

In relation to HIV infection, viral load has emerged as the dominant variable affecting infectivity (\( \beta \)). Blood concentration of HIV has been correlated with sexual transmission of the virus. Blood viral burden may act as a surrogate for HIV in genital secretions; however, at least one study found semen to be the most important vehicle for transmission of HIV, and the correlation between HIV in blood and HIV in semen appears to be poor. Co-infection with other STDs has also been associated with increased transmission of HIV. Recent analyses point to the importance of the population attributable fraction of HIV incidence associated with specific STD, during specific phases of STD epidemics, as a factor which affects the relative role played by other STDs in the spread of HIV. In this context, in recent studies that focus on HIV prevalence, risk attributable to HSV-2 infection has emerged as an important variable in HIV transmission.

Male circumcision is another factor associated with reduced acquisition of HIV infection. In one analysis circumcision markedly reduced HIV acquisition in high risk discordant couples, and reduced HIV transmission by HIV infected men with viral loads < 50 000 copies/ml. Methodological reviews of studies on circumcision and HIV transmission/acquisition suggest that confounding, effect modification, and estimates of population attributable fraction may underestimate the population level impact of circumcision. In most studies to date, adjustment for confounding by other risk factors has led to stronger associations. Circumcision effects vary across populations, perhaps because of effect modifiers such as age at circumcision, degree of circumcision, and the prevalence of other STDs. Effects appear to be stronger in populations at high risk of HIV and STDs. Epidemiological studies measure relative risks which reflect circumcision effects on male susceptibility at individual level; they do not measure circumcision effects on infectiousness or “herd effects.” Thus estimates of attributable fractions may underestimate the population level impact of circumcision.

The introduction of highly active antiretroviral therapy (HAART) has had considerable impact on the epidemiology of HIV and AIDS in resource-rich settings. HAART may decrease infectivity of infected individuals by reducing viral load. However, the population level effects of HAART are yet to be shown. While HAART has prolonged survival and decreased AIDS incidence, these beneficial direct effects may be counterbalanced by indirect mechanisms such as increased HIV prevalence owing to improved survival, increased unsafe sex owing to disinhibition of fear, increased STI incidence, and selective pressure for resistant viruses. The ability of antiretroviral treatment for HIV infected persons to prevent sexual transmission has not yet been demonstrated in humans or animal models. However, the biological plausibility of the approach is well established at the individual and dyad level. Overall population level impact of these treatments may only be properly evaluated through well executed community randomised trials.

The case of HAART provides a good example of how complicated the interactions among determinants of STD epidemics can be." The introduction of HAART functions as a determinant not only of \( \beta \) but also of \( C \) and \( D \). Moreover, the resulting changes in \( \beta \), \( C \), and \( D \) may further lead to additional changes in these determinants. The phase of the epidemic also affects the values of \( \beta \), \( C \), and \( D \). For example, as the AIDS epidemic got established in the United States, self reported condom use increased markedly. During recent syphilis outbreaks in different areas, as the outbreak matures, the duration of syphilis infection may get shorter because public health workers intensify their efforts to reach persons who are exposed to syphilis. Finally, it appears that in response to increasing prevalence of AIDS, people in many societies report decreasing their numbers of sexual partners.
Rate of contact between infected and susceptible individuals (C)

This particular determinant of STD epidemics has been long recognised to be of crucial importance. Many investigators have tried to capture the effects of this parameter by measuring the numbers of sex partners reported over different periods of time, such as one month, three months, a year, or a lifetime. Others have focused on the behaviours of core groups (persons with large numbers of partners who are interconnected with each other through sex links). The relative importance of numbers and patterns of sexual contacts of persons in the general population compared with the numbers and patterns of sexual contacts of core group members has remained controversial in the design of prevention programmes.

Interactions among the three immediate determinants of STD epidemics—β, C, and D—may influence whether sexual behaviours of the general population or the core groups are more important in determining the rate of spread of STI in populations. More specifically, for sexually transmitted infections with high infectivity (β) and short durations of infectiousness (D) such as gonorrhoea, the sexual behaviour patterns of core groups may be relatively more important in determining the rate of spread of STI. For infections with low infectivity and relatively long durations of infectiousness, such as HIV and HSV-2, the sexual behaviour patterns of the general population may be more important. In addition, the “dynamic topology” model postulates that during the initial growth and late low endemic phases of epidemics, while STIs are concentrated in core groups, the behaviours of core groups would be more important in determining the spread of infection.

Rate of contact between infected and susceptible individuals (C): the general population

Numbers of sex partners and rate of sex partner acquisition have been considered important variables that reflect the value of C for any population. Recently, timing of sex partnerships has emerged as an important dimension of such partnerships that influences the rate of contact between infected and susceptible persons (C). More specifically, the frequency of concurrent (overlapping in time) partnerships and the gap (duration of time between the end of one sex partnership and the beginning of the next sex partnership) have been found to be associated with the transmission of STDs. Mathematical models suggest that the prevalence of concurrency may be more important earlier in the history of epidemics in low prevalence situations, compared with higher prevalence situations in more mature epidemics; it is also more important in the dissemination of high infectivity—short duration infections than in the spread of low infectivity—long duration infections.

Concurrent partnerships occur often in both resource-rich and resource-poor settings. In one study, about 25% of a representative sample of reproductive age American women (15 to 44 years) reported having had concurrent partnerships over the past five years. Interestingly, in that study, 34% of women with a history of STD, compared with 25% of women with no STD history, reported concurrent sexual partnerships. Non-Hispanic blacks were significantly more likely to report concurrent partnerships (34%) than non-Hispanic whites (23%) and others (24%); women with lower family income were also significantly more likely to report concurrent partnerships than women with higher family income (30% and 22%, respectively). Proportions of men who report concurrent partnerships are even higher than those of women in the USA. In a recent study, 34% of men, compared with only 24% of women, reported concurrent sexual partnerships.

Even in the absence of concurrent partnerships, serial monogamy—in which the duration of time elapsed between last sex with one partner and first sex with the next partner, the gap—is shorter than the duration of infectiousness of any untreated infection would facilitate transmission of infection acquired from the first partner to the second partner. While not as effective in fueling the spread of STI as concurrent partnerships, short gaps may still play a large role in STD transmission dynamics because they apply to considerably greater proportions of men and women who practice serial monogamy. In a representative sample of reproductive age American women, 70% reported practising serial monogamy over the past five years. In this group, reported gaps were significantly shorter among women with STD history and among adolescents.

Specification of short and long gap lengths should depend on the duration of infectivity; as different sexually transmitted pathogens have different durations of infectivity, gaps that may be considered sufficiently long for short duration infections like gonorrhoea may be too short for long duration infections like herpes simplex virus or HIV. Duration of infection is also influenced by the existence and effectiveness of prevention programmes such as screening, and may vary across societies and perhaps across social classes within societies. Thus gap lengths which may contribute to the spread of a particular STI in poor rural areas of developing countries may be sufficiently long to limit the spread of the same STI in middle class urban areas of industrialised countries.

Mixing patterns are also an important factor to consider in describing the contact between infected and susceptible persons. Both epidemiological studies and mathematical models have shown that mixing patterns can fuel the spread of sexually transmitted infections in populations. Mathematical models often focus on the extent to which people of similar or dissimilar sexual activity classes mix with each other sexually. For STI that are incurable and not subject to suppression through antiviral treatment, sexual mixing across sexual activity classes may be a close approximation of contact between infected and susceptible individuals. For curable bacterial STI, and for STI that may be suppressed through antiviral treatment it may be important to consider both sexual activity and health care seeking in defining the groups involved in sexual mixing. Individuals who seek care promptly and who comply with treatment recommendations may not contribute to spreading infection, even if they are highly sexually active and have sexual contact with members of low sexual activity classes. One alternative that would incorporate the effects of both sexual activity and health care seeking may be to consider mixing between high and low STI prevalence populations. The relative role of sexual mixing patterns versus concurrency in facilitating STI spread had not been investigated until recently. One mathematical model revealed that in the population overall, mixing had a substantially greater impact on prevalence levels compared with concurrency. Whether this is generalisable to other values of the mixing and concurrency parameters, and to higher or lower prevalences of infection, is not yet known.

Patterns of sexual mixing, gaps between sexual partnerships, and frequency of concurrent sexual partnerships are all important determinants of the rate of contact between infected and susceptible persons in the general population. All of these variables show great heterogeneity across populations and over time, and they change in response to large scale societal changes. They are also likely to change in response to the evolution of STD epidemics as people adjust their behaviours to their environment. For the same reason, these variables may change in response to changes in β or D. Measurement and monitoring of these determinants of the contact rate between infected and susceptible persons in populations of interest would enable prevention programmes to design more effective interventions targeted at the appropriate population groups.
Rate of contact between infected and susceptible individuals (C): core groups

The role of core groups in maintaining STI in populations has been a focus of attention for STD researchers for many years. However, there is no current consensus in the STD field regarding the definition or relative importance of core groups. Various definitions have been used for this concept, some of which are categorical, such as “adolescents,” “sex workers” or “gay men.” Others are based on characteristics of networks that interconnect people with many partners to each other through sexual contacts. A recent attempt to compare alternative explanations for the persistence of STDs at low levels of prevalence revealed that, while heterogeneity in risk behaviour and long lived asymptomatic infections play a role in maintaining the prevalence of STDs, they fail to explain persistence of infections in local areas, suggesting that sex partner network approaches may be needed to explain the persistence of STI in local populations. A related mathematical modelling exercise explored the effects of network size on persistence of bacterial STD in populations. The network model was used to examine the relations among network size, network characteristics such as mixing patterns and levels of concurrency, and the persistence of infection over time. The results revealed that network size strongly influences persistence of infection. After 2000 days, 0.4% of networks of size 40 remained infected, 14% of size 1000, and 30% of size 15 000. Persistence of infection was also independently correlated with aspects of network structure, including the mixing pattern and the degree of concurrency. The network size and structure allow persistence of infection through the continued formation and dissolution of small groups of connected individuals with concurrent partners.

Most empirical work on core groups has focused on categorically defined study populations such as sex workers or gay men. Few studies have reported on study populations defined in terms of network connections. However, given the sexual lifestyles of sex workers, gay men, and adolescents, it may be quite appropriate to focus research attention on these groups as high sexual risk networks are considerably more likely to form around members of these categories than around other individuals. To the extent possible, it would be highly desirable to look at network connections around individual sex workers, gay men, or adolescents. The need for focusing research attention on clients of sex workers is glaring. However, studies of this population are notoriously difficult to conduct. Earlier studies have reported very high rates of condom use by female sex workers in the developed world, suggesting that sex worker populations in such contexts may not function as core groups for HIV infection. However, sex workers in other contexts, and even sex workers in developed countries, may not be well protected against other STI. In addition, recent large volume immigration of sex workers from other areas into developed countries may have resulted in decreased proportions of sex workers who use condoms or have health care access.

A desire to quantify and describe the distribution of risks for STI acquisition and STI transmission in core groups has led some researchers to quantify core group parameters such as size and structure. One analysis estimated the annual prevalence of full time equivalent sex workers in the United States for the period 1970–88 to be 23/100 000 and the mean number of male partners in the last 12 months per sex workers to be 694. A recent study estimated the number of men aged 15 to 49 years who visit sex workers in Cotonou to be 19 970 or 12.9% of the population, and calculated that, respectively, 1598, 14 618, and 459 low risk women in Cotonou would have been exposed to HIV, STIs, and simultaneously to both HIV and STIs through unprotected sex with male clients of sex workers during the past year.

In an assessment similar to that conducted in Cotonou, we estimated sex work related parameters in Saratov Oblast in Russia. In the Oblast region, Saratov, Engels, and Balakovo are the three epidemiologically, demographically, and economically distinct units and the social organisation of sex work varies across these sites. We collected data on number of sex workers, average number of nights a sex workers works a week, and the average number of clients they serve each night. The estimated numbers and rates of sex workers and sex worker–client contacts also varied across the sites. Such estimates provide a general understanding of the population structure underlying STI epidemics and may be helpful in guiding prevention activities.

Systematic data collection in a variety of societal contexts may help us understand the heterogeneity of population structures that underlie STD epidemics. A set of such variables may include: absolute numbers of sex workers, clients of sex workers, sex worker–client contacts, repeat clients per sex worker, and repeat visits to specific sex workers per client. These variables may be assessed in terms relative to the size of the population base as well. The relative measures would include number of sex workers, number of clients of sex workers, and number of sex worker–client contacts per 100 000 population. These parameters are based on numbers of sex workers and their clients; they do not take into consideration the fact that some sex workers remain on the job practically all waking hours, while others may work only one evening a week, and still others may do sex work on an “as needed” basis. Four additional variables may help us understand the time dimension of sex work and incorporate it into the estimates for sex workers and sex worker–client contacts. Absolute parameters that may be considered in this context include available sex work hours per month and hours of sex worker–client contact per month. If these variables were to be assessed in terms relative to the size of the population base, they would be: available sex work hours per month per 100 000 population; and hours of sex worker–client contact per month per 100 000 population.

Measures of concentration of sex work may also be important. What proportion of sex workers provide services to what proportion of clients per month? What proportion of sex workers provide what proportion of hours of sex worker–client contact per month? Highly concentrated sex work, both in terms of the number of clients served and of the number of hours of sex worker–client contact, may have a much greater impact on the spread of STD than widely distributed sex work loads. Measures of concentration such as Gini coefficients and Lorenz curves may be applicable in this context. These variables, while they may not directly inform network aspects of core groups, may provide similar information without the great complexities of network data collection. Information on these variables has been rare in published reports up to now. However, a two pronged approach, where mathematical modelling efforts focus on the potential role and impact of these variables in different settings—marked by prevalence levels, phases of epidemics, and demographic and socioeconomic context—and empirical work focused on assessments of these variables in a variety of contexts, may help enhance our understanding of the evolution of these epidemics greatly.

Focus on these issues led to various empirical and methodological questions. One of these relates to the size of the core groups; does it remain constant or change, perhaps in response to economic pressures? Bloor et al suggest that the size of the sex worker population that an area can economically sustain may be fixed. This suggestion may be overly simple. Fees for sex work are quite variable and can change drastically over time; this was observed following the 1998 economic crisis in Moscow. Such changes in price may dramatically alter the relation between the size of the sex worker group and the environment, thereby allowing the size of the group to change. Moreover, factors external to the sex worker–client system, such as relevant technologies and geographical/social mobility, may change the social organisation of sex workers, allowing the size of the sex worker group...
to change. The introduction of cell phones and the internet into the social organisation of sex work has greatly increased the mobility of sex workers and their ability to service an enlarged group of clients in London (Ward H, personal communication). It is not known how the size of sex worker groups has been affected by these changes; it is certain, however, that the relation between the size of the sex worker group and the so called environment is affected by these changes. Information on the factors that limit the size of the sex worker population is limited. It is likely that the drastic demographic and technological changes taking place have an impact on these factors. In the Saratov Oblast,1 the factor that limited the volume of sex worker–client contacts was the number of visits to different sex work venues. Demand for sex was so high that the sex worker group was unable to meet the demand for escort services. The volume of client demand for sex worker services also exceeded the volume available slots within these services. The volume of client demand for sex worker services was limited by the availability of escort services. The factors that limit the size of the sex worker population in particular, and the size of core groups in general, probably vary across social contexts. It would be helpful to know what these factors are and how they vary.

Population turnover in sex worker groups in particular, and core groups in general, is another related but distinct issue. Within defined time periods, to what extent does membership of these groups change, through new individuals entering and leaving? What is the rate at which exited individuals re-enter core groups? What is the impact of such turnover on STD epidemics? Do new entries, repeat entries, and exits have a differential impact on STD rates? And what are the societal factors that influence these parameters? A two pronged approach, with both mathematical modelling and empirical research focusing on these issues, may lead to new findings and help move the field forward. Measurement of individual based variables, length and timing of sex work careers, and length and timing of careers as a sex worker client may help assess population turnover. Several studies have shown that the definition of core groups should not be limited to easily recognised categorical groups such as sex workers. For example, studies in rural North Carolina in the United States52 and rural areas of Zimbabwe53 identify rural women at increased risk of STDs, including AIDS, whose general rate of partner change may be lower than that in the sex worker group in urban areas. Particularly in the light of changing societal and economic conditions, this may be an important point to consider in studies of core groups.

Many of the concepts discussed above, including the absolute and relative size of core groups and the absolute and relative volume of contact between core groups and the general population, are already in the STD literature.53–54 However, empirical evidence regarding these concepts across societal contexts is limited, and the impact of these variables on STD incidence/prevalence or the evolution of STD epidemics is unexplored.

Overlap among risk behaviours has been well established. For example, the overlap between drug use and sex work appears to be universal.54–55 In addition to some groups of female sex workers, highly sexually active gay men may also constitute cores for overlapping sexual and drug use related risks. In one study, 95% of men reported drug use (excluding Viagra) during the most recent circuit party weekend they attended. In this sample, 12% of the men reported using Viagra during circuit party weekends.56 Another study conducted in Houston, Texas,57 reported highly overlapping risk behaviours (including drug use and homosexual and heterosexual contact with large numbers of partners) among members of a densely interconnected core group of men who have sex with men in exchange for money. Some of these men were highly transitory, moving into and out of the core group frequently, averaging about three months in the core network. In addition, this group of men was highly mobile, moving across cities as well as within the city. High levels of mobility have also been noted for female sex workers in Africa and Asia. During the 1990s major influxes of sex workers have been observed all over Europe. Some of these migration streams originate in the former Soviet Union and eastern Europe, others in Africa and Asia. For example, the proportion of women with non-UK origins was 24% in a sex workers cohort recruited from 1985 to 1991 in London, this proportion increased to 57% in the cohort recruited between 1995 and 2000. In 2000, 34% of sex workers were from eastern Europe and the former Soviet Union.58 This type of migration may have important effects on the evolution of STD epidemics. Often immigrant women lack adequate access to preventive and health care services; their levels of condom use may be relatively low and the duration of their infections may be longer than that for the resident sex worker populations. In low STD prevalence settings, these groups may be particularly important as they may constitute an effective mechanism for the reintroduction and persistence of infection.

Technological, social, economic, and demographic developments may result in important and rapid changes in the size, structure, distribution, and function of core groups. For example, in Moscow before the August 1998 economic crisis the number of female sex workers was estimated at 15–30 000 following the crisis this number increased to 30–90 000. Similarly, in Indonesia between 1997 and 1998 (following the monetary crisis), the percentage of female sex workers who were less than 20 years of age increased by 38% and 125% in Jakarta and Surabaya, respectively. The percentage of female sex workers with less than 12 months’ experience increased by 28% and 130% during the same period in these two cities.59 Proportions of women younger than a given age, and with less than a specified amount of experience, may provide an easy way to assess indicators for monitoring recruitment into core groups.

Assessing and monitoring removal from core groups may be as important as monitoring recruitment into core groups, but for different reasons. A recent analysis of AIDS mortality and syphilis rates in the USA revealed that AIDS deaths, by removing core group members, particularly among men having sex with men, may have contributed to declining syphilis rates in some states between 1983 and 1997.60 However, after HAART the improvements in AIDS survival in the USA may have contributed to the recent outbreaks of syphilis in homosexual men.

Finally, recent developments in communication technology have drastically changed the social organisation of sex work in places. The advent of cell phones, pagers, and the internet has increased sex workers’ freedom to move about (as they no longer have to wait in a particular place), and has expanded the sphere of their potential clients. However, these developments may have increased the competition from other sex workers at the same time. One recent report61 suggested that the Dominican Republic has become one of the biggest sex tourism destinations in the world, mostly as a result of internet sites that extol the country as a “single man’s paradise.” Owners of websites which include information on the sex workers and the services they provide encourage as many pimps as possible to advertise on their sites for a fee. These developments are global and are drastically changing the mosaic of the sex industry (Ward H, personal communication). Anecdotal evidence suggests that this technological evolution has greatly increased the size of sex worker core groups and multiplied manyfold the numbers and geographical reach of interconnections among sex workers and their clients.

**Duration of infectiousness (D)**

The duration of infectiousness, the third immediate determinant of the rate of spread of STI in populations, may be easier
to modify through public health interventions, compared to β and C. For curable bacterial STD, timely diagnosis and appropriate treatment have been the most time honoured interventions, and screening asymptomatic men and women—have a major impact on the duration of infectiousness, thereby reducing rates of STI. In one study, provision of monthly presumptive treatment to women at high risk in a mining community significantly reduced STD rates among miners between 1996 and 2000. The frequency of presumptive treatment was reduced from monthly to quarterly in 2000; however, STD rates remained consistently low. Screening asymptomatic women for chlamydial infection is a well established programme priority in many developed countries. Extension of such screening to non-traditional settings and to cover asymptomatic men is currently being implemented.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.

Programmatic interventions—including partner notification, syndromic management, periodic presumptive treatment, and screening asymptomatic men and women—have a major impact on the duration of infectiousness. The person–time-of-infectiousness model suggests several major types of delay. Diagnostic delays cover time elapsed from patient-provider contact to clinical or laboratory diagnosis, and in the case of laboratory diagnoses, time elapsed from laboratory diagnosis to the patient receiving the diagnosis. Treatment delays include time elapsed from receipt of diagnosis to receipt of drug treatment; from receipt of drug treatment to completion of drug treatment, and finally to cure or suppression of infection. In the USA, studies show that health care seeking delays range between 5.8 and 9.6 days for adolescents. One third of women with symptoms delay seeking care by at least one week, and among women with pelvic inflammatory disease, over 65% do not complete treatment.
the internet), and transportation, and consequently, shared experiences. In everyday life, this contraction in the spatial dimension is often experienced as “the small world.” In the temporal dimension, globalisation refers to decreases in the actual and perceived time required or allowed for human activity to occur. In everyday life the contraction in the temporal dimension is experienced as an accelerated time frame. In the cognitive dimension, globalisation includes the globalised production of knowledge, ideas, beliefs, and values. In everyday experience, these changes reflect themselves in homogenising cognitive processes.

One chronology of stages of globalisation in human history was provided by Robertson in 1992 (fig 3). Globalisation has been related to the patterns of evolution of other infectious diseases, for example cholera. No specific studies have been conducted on the relation between globalisation and the evolution of STD epidemics; however, patterns of association are clearly suggested. The contracted spatial and temporal dimensions allow people located at great distances from each other to connect over the electronic media and get together rapidly, through accelerated time frames. The changes in the sex work industry and the gay men’s circuit parties discussed earlier are examples of the impact of globalisation on one of the determinants of STD epidemics (C). The net effect of globalisation on the evolution of STD epidemics needs to be assessed in future research. One hypothesis would be that the evolution of STD epidemics will be expanded spatially and contracted temporally as it is affected by globalisation, just like many other evolutionary processes.

### Societal Change and the Evolution of STD Epidemics

Human societies have evolved from the hunting and gathering society, through the agricultural revolution and the industrial revolution. The more recent decades have witnessed the information revolution and the establishment of the global village, together with increases in population size, concentration of population in space (urbanisation), volume of migration, and the extent of inequality. At the same time major changes have taken place in the functions of the family and the status of women. How would these global societal changes affect the evolution of STD epidemics through specified phases?

Some alternative ways in which the evolution of STD epidemics may be affected by ongoing societal changes can be suggested. First, the pattern of evolving STD epidemics may remain the same, but movement through the epidemic phases may happen faster, collapsing the time it takes for STD epidemics to evolve. Second, the pattern of the evolution of STD epidemics may change as a result of the rate of exposure between infected and susceptible individuals (C) increasing continuously at accelerating rates owing to increased inequality and globalisation; and duration of infectiousness (D) increasing continuously as a result of societal collapse in an increasing number of countries. Third, globalisation may result in spread, maintenance, and dead-end networks becoming connected with each other globally. Such connection may have several consequences:

- local epidemics may become interdependent, making their identification difficult;
- spread networks may get concentrated among the poor locally and get connected to each other globally;
- dead end networks may get concentrated among the rich and get connected globally;
- in spread and dead end networks, the rich and the poor may get further disconnected from each other.

It is difficult to know exactly how these predicted changes may affect STD epidemics; however, systematic observation of these predicted changes will greatly enhance our understanding of STD transmission dynamics in a changing world.

In more general terms, the interactions among globalisation, inequality, and STD transmission dynamics may be conceptualised in terms of increasing inequality, increasing globalisation, and increasing proportions of the population living in post-conflict/post-collapse societies, leading to simultaneous increases in (C) (owing to increases in the absolute and relative size of core groups and in prevalence of concurrency) and in (D) (owing to collapse of health systems, decreasing health care seeking, declining health service provision, and resulting delays in diagnosis and treatment), as well as increased interconnectedness between core groups across local areas and increased interconnectedness between general populations across local areas (fig 4). The main issues regarding the impact of globalisation and increased inequality on STD prevention and control efforts may be reduced to two major questions. First, globalisation and increased inequality will lead to interconnected core populations across diverse geographical areas and interconnected general populations across diverse geographical areas, mixing over shorter time periods, which will probably lead towards globalised epidemics. In this context, what is the ability of the local public health systems to deal with global epidemics? Second, as a result of

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time scale</th>
<th>Spatial centre</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germinal</td>
<td>Circa 1500–1850s</td>
<td>Europe</td>
<td>Growth of national community, accentuation of the concept of the individual, spread of Gregorian calendar</td>
</tr>
<tr>
<td>Incipient</td>
<td>Circa 1850–1870s</td>
<td>Mainly Europe</td>
<td>Shift towards homogeneous unity of the state, formalised international relations</td>
</tr>
<tr>
<td>Take-off</td>
<td>Circa 1870–1920s</td>
<td>Increasingly Global</td>
<td>Inclusion of non-European states into international society, WWI, League of Nations</td>
</tr>
<tr>
<td>Struggle for Hegemony</td>
<td>Circa 1920–1965</td>
<td>Global</td>
<td>Wars/disputes about the shape of the globalisation process, atomic bomb, UN</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>Circa 1965–1990s</td>
<td>Global</td>
<td>Inclusion of Third World, moon landing, end of Cold War, environment, HIV/AIDS</td>
</tr>
</tbody>
</table>
globalisation and increased inequality, the gap between high and low socioeconomic classes will widen, which will mean that the gap between those who control health systems and those who are infected also widens. Under such circumstances, how do power elites deal with STDs among populations increasingly distanced from themselves socially, economically, and perhaps politically?

A reasonable projection of probable changes in the pattern of disease is essential to an understanding of the evolution of disease control priorities. In predicting changes in pattern of disease, Mosley and colleagues considered the effects of the demographic transition, the epidemiological transition, changing risk environments, and the widening gap in health problems and health needs across social and economic classes; together these changes were termed “the health transition.” In the health transition model, the demographic transition leads to declining infectious disease mortality, declining fertility, and an aging population; the epidemiological transition leads to the emergence and increase of chronic and non-communicable diseases; and the protracted-polarised epidemiological transition leads to the persistence or re-emergence of communicable diseases as a result of economic recession and increasing inequality. The health transition model may be adopted to the evolution of STD epidemics (fig 5). The STD health transition model would suggest that future decades may witness increases in viral STD and the persistence and re-emergence of bacterial STD.

CONCLUSIONS

Effective planning and evaluation are essential to the management of successful STD prevention programmes, particularly in limited resource environments. Enhanced surveillance is crucial for planning and evaluation. Monitoring the proximate and underlying determinants of STD epidemics may lead STD prevention programmes to a better understanding of the factors that influence trends in morbidity and to better predictions of future changes in these factors. A conceptually grounded, phase specific approach to enhanced surveillance of determinants of STD epidemics may be the

Figure 4 Globalisation, inequality, and sexually transmitted disease (STD) transmission dynamics. Dx, diagnosis; Rx, treatment.

Figure 5 Sexually transmitted disease (STD) health transition.


