Populations, pathogens, and epidemic phases: closing the gap between theory and practice in the prevention of sexually transmitted diseases

J F Blanchard

The phase specific model for the prevention and control of sexually transmitted infections (STI) offers new insights into the strategic planning of programmes. The model illustrates the importance of modifying the focus of prevention and control activities to different subpopulations as the epidemic evolves over time. However, the practical application of phase specific approaches will depend on an understanding of the variability and determinants in the trajectory by which STI epidemics progress through epidemic phases. This paper draws on empirical observations from diverse populations to explore the influence of sexual behaviour patterns in populations, the biological characteristics of STI pathogens, and the population-pathogen interactions in relation to epidemic trajectories. In addition, various approaches to the determination of epidemic phase are presented.

In the past two decades, there has been a substantial increase in theoretical knowledge about the transmission dynamics of sexually transmitted infections (STI). This has largely been motivated by the global epidemic of HIV and the vigorous efforts to reduce the burden of other STIs, and has been facilitated by the emergence of new methods and technologies for data analysis and mathematical modelling. Arising from this research is an increasing appreciation for the importance of the size and nature of core groups in the spread and maintenance of STIs within populations.

Recently, these concepts have been enriched by mathematical modelling and empirical research showing the importance of the structures of sexual networks in a population in determining the rapidity of spread and, ultimately, the prevalence of STIs that can be sustained within a population. This improved understanding of the transmission dynamics of STI could have far reaching implications for the design of highly strategic prevention programmes.

The core group concept suggests that there might be an epidemiological “bulls eye” for prevention programmes, and increasingly public health practitioners are exploring the possibility of intervening with prevention programmes through specific sexual networks. However, as Wasserheit and Aral have pointed out, the distribution of STIs within a population is not static. Rather, they propose that over time STI epidemics evolve through different phases, which are characterised by changing patterns in the distribution and transmission of STI pathogens within and between subpopulations. This implies that the relative importance of core groups to the spread and maintenance of STIs within a population will change over time.

This elegant formulation of the dynamic interplay of STI epidemics and programmatic responses over time represents an important conceptual step forward. Ideally, it should also have considerable practical implications by providing programme planners with the ability to develop a “road map” for tailoring prevention programmes such that they are appropriate to the phase of the epidemic. However, as with any new theoretical framework, some further elaboration of the phase specific model is necessary to enhance its relevance to the circumstances encountered by programme planners and implementers at ground level. In this regard, an important theoretical and practical issue for the phase specific model relates to understanding the variability and determinants of the trajectory of STI epidemics over time.

Although the phase specific model describes a generally stereotypical progression over time in terms of the incidence and prevalence of disease and the subpopulations in which the STI is being transmitted, Wasserheit and Aral point out that the uncontrolled course of an epidemic will vary depending on the patterns of sexual behaviour and networks in the population, and the biological properties of the pathogen. Thus some epidemics are likely to progress rapidly through the early epidemic phases to reach a “hyperendemic” phase in terms of disease prevalence and population distribution, whereas others may progress slowly or become “arrested” at an earlier epidemic stage. From a practical perspective, developing methods for predicting and monitoring the epidemic trajectory over time will have important implications for prevention programmes. For example, in circumstances where a rapid expansion of the epidemic from relatively small high risk core groups to the more general population is anticipated, the programmatic response should be planned accordingly with the early implementation of activities that are directed to the general population. Conversely, if the epidemic transmission dynamics are not likely to progress beyond a dependence on a relatively small core group within the population, the programmatic response should maintain a steady
focus on those subpopulations. However, the theoretical and empirical bases for practical approaches to predict the epidemic trajectory and assess the epidemic phase are underdeveloped.

In October 2000 a scientific meeting was held in Rome to bring together researchers and programme planners from around the globe to present and discuss empirical observations of diverse STI epidemics in an effort to gain a better understanding of the implications of the phase specific model. The purpose for this paper is to draw upon these observations to elaborate on the phase specific model with a focus on two issues: first, approaches to categorising and predicting different trajectories by which STI epidemics will progress through the epidemic phases, and second, methods for identifying the epidemic phase.

CATEGORISING AND FORECASTING EPIDEMIC TRAJECTORIES

As Wasserheit and Aral have pointed out, the progression of an STI epidemic can be viewed from two main perspectives. The epidemiological perspective focuses on the incidence and prevalence of disease in the population. Accordingly, the general progression of an uncontrolled epidemic will be an initial rise in incidence and prevalence (“growth” phase), followed by stabilisation in incidence and a plateau in prevalence (“hyperendemic” phase). The institution of an effective prevention programme will reduce the incidence and prevalence and push the epidemic into a “decline” phase, which may move all the way to disease elimination but more often results in a fluctuant epidemic state of lower prevalence with intermittent spikes of incidence and prevalence.

An alternative perspective for categorising the progression of an epidemic over time focuses on the transmission dynamics within the host population. This approach has greater importance for prevention programmes, as it specifies the subpopulations within and between which the STI is being primarily transmitted and thus provides guidance for focusing prevention activities. In this regard, Wasserheit and Aral have suggested that early in an epidemic (that is, during the growth phase), an STI pathogen is likely to be transmitted within and from high risk “spread networks.” As the epidemic progresses the pathogen spreads more into lower risk populations where there are “maintenance networks,” and ultimately reaches equilibrium with the host population (during the hyperendemic phase). Thereafter, the course of the epidemic (the “controlled history”) is largely determined by the response of the population with respect to changes in sexual behaviour patterns and specific prevention programmes.

Clearly, STI epidemics will have different trajectories as they progress through the epidemic phases. The focus of this paper is on the uncontrolled epidemic because, once prevention activities are initiated, the subsequent course of the epidemic is highly dependent on the quality and coverage of the control programme. In considering the categorisation of these trajectories, it is important to clarify the strategic goal of a phase specific approach. It is not the selection of the specific content of prevention programmes, as these may differ according to local circumstances, but rather it is the determination of how best to focus prevention programmes to subpopulations—defined by sexual behaviours and sexual networks—at various phases of the epidemic. With specific reference to categorising and predicting different epidemic trajectories, the central questions are the extent to which and how quickly there will be a shifting of the transmission dynamics from a dependence on high risk core groups to more widespread transmission in the general population. Thus it is important to know whether and when there are likely to be “inflection points” in the epidemic, which are characterised by a shift in the relative importance of different subpopulations for the continued growth and maintenance of the epidemic.

It should be noted that the term “core group” in this discussion is not based on a theoretical construct, but is instead meant to have practical programmatic relevance in so far as it refers to relatively small subpopulations which can be identified for targeting of prevention programmes. On the basis of these strategic considerations, at least three broad patterns or trajectories of epidemic progression can be envisioned (fig 1).

The first trajectory (trajectory I) is where the epidemic begins in a small core group in which it expands fairly rapidly, but never escapes much beyond that subpopulation in terms of its ongoing transmission. Although lower risk members of the population will also be infected, the hallmark of this pattern is the perpetual dependence on relatively small core groups. Thus the evolution of the epidemic is “arrested” in relation to the phases described by Wasserheit and Aral, as it never progresses to the point where there is important propagation of the epidemic beyond the initial core group into which it was introduced.

A second trajectory (trajectory II) would follow closely the pattern described by Wasserheit and Aral. In this pattern, the epidemic is initially introduced and expands within core groups, but subsequently seeds lower risk populations where the sexual behaviours and networks are sufficient to maintain the epidemic but are not conducive to continued epidemic expansion.

The third trajectory (trajectory III) would entail the rapid seeding of a larger proportion of the lower risk population, where the STI pathogen would subsequently maintain a substantial foothold and could continue propagation independently of easily identifiable core groups.

Although this classification of epidemic trajectories is unrefined, programme planners would benefit from better methods for predicting which general pattern an STI epidemic will follow. Such methods will rely on further understanding of how the epidemic trajectory will be influenced by the patterns of sexual behaviours and networks in the population, the biological characteristics of the STI pathogen, and the consequent pathogen–population interactions.

---

**Figure 1** A general classification scheme for epidemic trajectories, adapted from Wasserheit and Aral. Pattern I illustrates epidemics which begin and expand in relatively small subpopulations or core groups (dark circles), and ongoing transmission continues to be directly dependent on these core groups throughout the epidemic. Pattern II illustrates an epidemic trajectory where after the initial seeding and expansion of the epidemic in core groups, there is an expansion to lower risk involve lower risk segments of the population (represented by light circles), though core groups continue to play an important role in transmission. Pattern III represents epidemics where there is rapid expansion beyond definable core groups such that continued epidemic maintenance and growth is somewhat independent of easily identifiable core groups.
Population sexual behaviour and network characteristics

There are various ways of classifying the patterns of sexual behaviour and networks in a population. General classification schemes have used terminology such as “core” and “non-core” groups to denote those members of the population with high and low rates of sexual partner change. A substantial proportion of most populations will have very low rates of partner change, with many having only a small number of lifetime sexual partners. Thus populations can be described on the basis of the prevalence of persons with high rates of partner change. However, as Garnett has pointed out, for STI transmission it is not just the number of partners that matters, but also the duration of partnerships. So it is also important to know the proportion of the population with multiple, long term partnerships, either concurrently or serially. This variable is important as the duration of partnerships influences the probability of transmission per partnership.

Thus the general patterns of sexual behaviour in a population can be crudely classified according to the proportion of the population with a large number of short term partnerships (such as commercial sex workers) and the proportion with many partnerships of long duration, either concurrently or serially. On this basis, an assessment of the empirical data and anecdotal reports of sexual behaviours and networks in different populations described in this volume yields several discernable patterns.

Although quantitative data are not adequate for precise classification, fig 2 provides a summary of these patterns. At one end of the spectrum (pattern A) are populations such as England and Wales, where only a small proportion of the population has many short duration partnerships or many long duration partnerships. At the other end of the spectrum (pattern D) are populations in which the prevalence of both types of sexual partnership is high. These include Zimbabwe and Nairobi, Kenya, where female sex work is common and where a sizable proportion of the population also has multiple sexual partnerships of longer duration, either concurrently or serially. There are also populations where the pattern is mixed. Hawkes and Santhya describe the situation in India where commercial sex work is extensive, citing studies indicating that 8–29% of men have paid for sex. However, concurrent or serial partnerships of longer duration appear to be generally much less common in India owing to the prevailing cultural norms. Similarly, Rekart reports that in Ho Chi Minh City, Vietnam, commercial sex is quite common, whereas the frequency of multiple longer duration partnerships is relatively low. Lowndes and colleagues report that in Cotonou, Benin, the proportion of the population that is involved in commercial sex work is high, with approximately 13% of men aged 15 to 49 years visiting a female sex worker each year. However, the prevalence of multiple long duration partnerships appears to be lower, especially among women. Thus in these populations, the pattern (pattern B) is one where a relatively large proportion of the population has many short duration relationships, but few have multiple long duration relationships. In contrast, the papers by Changalucha et al and Boerma and colleagues describe a pattern (pattern C) in Mwanza, Tanzania, where commercial sex work with large numbers of clients is not as common, but a relatively high proportion of both men and women have several concurrent or serial longer term sexual partnerships through polygamous relationships or divorce and remarriage.

Transmission properties of STI pathogens

The ecological success of an STI pathogen depends largely on two characteristics—its transmission efficiency per contact (infectiousness) and the duration of infectiousness (β and D), respectively, according to the formula for the basic reproductive rate of an STI as described by May and Anderson. These parameters have been estimated for several different pathogens. Hence, as Garnett points out, STI pathogens can be roughly categorised on the basis of these two characteristics (fig 3). Pathogens such as HIV and herpes simplex virus 2 (HSV-2) are of relatively low infectiousness, but are of long duration (that is, low β, high D). At the other end of the spectrum are pathogens such as Haemophilus ducreyi that are highly infectious but of shorter duration (high β, low D). Pathogens that are of intermediate infectiousness and duration include Chlamydia trachomatis, Neisseria gonorrhoeae, and Treponema pallidum.

Population–pathogen interactions: implications for epidemic trajectory

The pattern or trajectory by which an STI epidemic will evolve through epidemic phases will differ for different types of
population–pathogen interactions (table 1). For example, in settings where only a small proportion of the population have very high contact rates or many long duration partnerships (that is, pattern A in fig 2), it is likely that for pathogens of either short duration (such as *H. ducreyi*) or low infectiousness (HIV) the epidemic will be continuously highly dependent on a small core group in the population. Thus STI epidemics in these circumstances are likely to follow a truncated trajectory (trajectory I in fig 1), and would not require substantial shifts in the focus of prevention programmes beyond those subpopulations. In such populations only those pathogens that are highly infectious and of reasonably long duration (for example, *C. trachomatis*) will be likely to follow an epidemic pattern where there is a phase of substantial transmission that is not directly linked to core group transmission (trajectory II in fig 1). At the other end of the spectrum, for populations where there are relatively large subpopulations with high contact rates, interconnected with a large segment of the population with multiple, longer duration partnerships, the trajectory would be such that for most pathogens there will be a rapid expansion of the epidemic beyond the core group (trajectory III).

Using these broad generalisations, it is possible to hypothesise about the epidemic trajectory and implications for control strategies for specific situations described in this volume. For example, for populations such as in Manitoba or England and Wales, *C. trachomatis* appears to have followed a trajectory that resulted in a substantial amount of transmission occurring outside of easily defined core groups (that is, trajectory II). However, there is little evidence, at least from England and Wales, that sexual transmission of HIV has escaped beyond relatively small subpopulations, thus indicating that it has followed a truncated trajectory in terms of population distribution (trajectory I). In contrast, in Zimbabwe, where there are relatively large segments of the population with high partner change rates including many long duration partnerships, HIV appears to have spread quite rapidly among the general population (trajectory III). Similarly, in Mwanza, Tanzania, where multiple, long duration partnerships are common but female sex work is less so, the HIV epidemic is not directly linked to core group transmission (trajectory II). Hence, as it is uncommon for women not involved in sex work to have multiple concurrent or serial longer term relationships, Lowndes and colleagues recommend that even 15 or more years into the epidemic a continued strategy that focuses on reducing HIV transmission to and from core groups comprised of female sex workers and their clients is warranted.

Another example of the difference in epidemic patterns based on population sexual network structures is offered by a mathematical modelling exercise by Nagelkerke and colleagues. Using a compartmental model, they compared the population level impact of various prevention strategies on HIV prevalence in two populations: Botswana and India. In comparing the two populations, the crucial assumptions were: (1) the proportion of men who will have a sexual contact with a female sex worker is fourfold higher in Botswana, and (2) the number of “low risk” persons of the opposite sex infected by an HIV infected individual is twice as high in Botswana (this parameter is closely related to the frequency of multiple, longer duration partnerships). Thus the model implicitly assumed that the size of population with high contact rates and multiple longer duration partnerships is substantially greater in Botswana than in India. It also assumed that the prevention interventions were in effect from when the epidemic had reached a prevalence among adults of approximately 2.5% in India and 28% in Botswana. The mathematical models showed that for the Indian scenario a single prevention strategy—begun when the population prevalence was 2.5%—which reduced the proportion of sexual contacts between female sex workers and their clients who were unprotected by condom use from 67% to 25% would be sufficient to arrest the HIV epidemic and eventually drive it towards extinction. However, in a setting such as that assumed for Botswana, such a targeted intervention would only stabilise the epidemic at the current prevalence, whereas an intervention that reduced HIV transmission more generally in the population (by decreasing the overall burden of treatable STIs) would reduce the prevalence of HIV substantially. Unresolved from these analyses is the question of whether initiating the different prevention strategies at different stages of the epidemic would have altered their relative impact.

**IDENTIFYING THE EPIDEMIC PHASE**

The various approaches that have been used to identify the epidemic phase can be generally grouped into two categories: the “pathogen perspective” and the “population perspective” (fig 4).

The first approach (“pathogen perspective”) is based on the relative success of the pathogen in propagating and maintaining an epidemic, which is assessed by examining the descriptive epidemiology of the STI pathogen. Thus the phase of the STI epidemic is determined by either the current level of the

<table>
<thead>
<tr>
<th>Pathogen characteristics</th>
<th>Population prevalence of sexual behaviours patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>High infectiousness, short duration (eg, <em>H. ducreyi</em>)</td>
<td>Large high contact, long duration (eg, Mwanza)</td>
</tr>
<tr>
<td>Intermediate infectiousness and duration (eg, <em>C. trachomatis</em>)</td>
<td>Large high contact, long duration (eg, Botswana)</td>
</tr>
<tr>
<td>Low infectiousness, long duration (eg, HIV)</td>
<td>Large high contact, large long duration (eg, Zimbabwe)</td>
</tr>
</tbody>
</table>

*Epidemic trajectories listed in the body of the table are those illustrated in fig 1.*
time trend in the prevalence or incidence of the disease. For example, populations in which there is currently a high prevalence of disease are usually classified as being in a hyperendemic phase. In populations where the disease incidence and prevalence are increasing the epidemic is judged to be in a growth phase, and where the disease incidence or prevalence is declining the epidemic is said to be in decline.

While this approach may provide an accurate description of the epidemic phase from the pathogen perspective (that is, the relative propagative success of the pathogen), it has important limitations. First, as the transmission dynamics for different populations vary widely it is difficult to establish clear guidelines for assigning the epidemic phase, especially when the criterion is the absolute prevalence or incidence of disease. For example, Moses and colleagues have used data based on the monitoring of STI prevalence among antenatal women to describe a transition from a hyperendemic phase to a decline phase for chlamydia (33% to 17%). Low has indicated that chlamydia is in a growth or hyperendemic phase in specific wards of London, with rates of reported cases among women of approximately 300 per 100 000 annually and increasing. While the time trends in these two populations appear to justify the phase categorisation, the absolute rates of disease are hard to reconcile into the same category.

The UNAIDS and WHO have suggested a classification scheme wherein numerical proxies are used to describe three different HIV epidemic phases (“epidemic states” by their terminology). Low level” epidemics are those where the HIV prevalence has not yet reached 5% in any defined high risk subpopulation (for example, female sex workers, injection drug users); “concentrated” epidemics are those where the HIV prevalence is consistently over 5% in at least one defined subpopulation but below 1% among urban pregnant women; and “generalised” epidemics are those where the HIV prevalence is consistently over 1% among pregnant women. Although these epidemiological approaches to identifying epidemic phase offer the advantage of being practical and accessible in most settings, they do not necessarily assist in determining when a shift in prevention strategies is warranted. For example, the HIV epidemic situation described by Lowndes et al in Cotonou would clearly fall into the “generalised” epidemic category of the UNAIDS/WHO framework, but they have shown that the epidemic is still driven largely by female sex work, thus justifying a continuation of a prevention strategy that is focused to preventing transmission in that setting.

For a better guide to the selection of phase appropriate prevention strategies, methods that identify important shifts in the transmission dynamics are required. Thus another set of methodologies has emerged that attempt to assess the epidemic phase by determining how the pathogen is distributed and transmitted within a population (the “population perspective”). In that regard, several approaches have been used. One approach is the assessment of the geographical “concentration” of STIs over time. The underlying assumption for this method is that certain geographical areas, usually characterised by having populations that are of low socioeconomic status and marginalised in other ways, will contain a greater proportion of STI core group subpopulations. Therefore, in situations where there is little geographical concentration of an STI, it is assumed that the epidemic is more likely to be at a hyperendemic phase, as there is an assumption that the STI is widely distributed into the maintenance networks in the population. In contrast, where an STI is highly concentrated geographically, it is assumed that the epidemic is at an early stage or in a late decline phase, as such a pattern would indicate that the STI is concentrated more in the core groups or spread networks.

This approach has some appeal because it does not require a detailed specification of the spread and maintenance networks and can be accomplished with routine surveillance data. It also has important shortcomings. First, the method assumes that surveillance systems accurately reflect the distribution of STIs in the population, and that there are no important temporal changes in the sensitivity and specificity of surveillance case definitions and methods. Second, there is no widely accepted measure of geographical concentration that is theoretically robust and has clear practical applicability.

Another approach to identifying epidemic phase from the population perspective is through the analysis of sexual networks. Potterat and his colleagues have analysed sexual network structures based on contact tracing data as a method for identifying epidemic phases for chlamydia. This method appears to get closer to identifying the actual transmission dynamics, and therefore might offer more useful data for strategic decision making. However, there are still problems with this approach that need to be resolved. From a practical perspective, mapping out sexual networks requires an adequate infrastructure for contact tracing to obtain an accurate snapshot of sexual networks through which STIs are propagating. Furthermore, such an approach may not be suitable for STIs of long duration such as HIV, as the appropriate time frame for identifying sexual contacts will be much longer.

A third approach for determining epidemic phase is through mapping out “transmission chains.” Although they have not used it to identify the epidemic phase explicitly, Lowndes and colleagues have shown this approach for the population of Cotonou, where they have employed data from various sources to determine the extent to which HIV transmission is dependent on female sex work. If, instead, such an exercise revealed that there was substantial transmission occurring apart from definable core groups, that might be a useful indicator that the epidemic had progressed beyond dependency on spread networks for transmission, with consequent implications for the focusing of control programmes. Thus that approach could be of substantial practical value for guiding phase specific strategies. However, for such an approach to be broadly applicable, rapid methods for mapping transmission chains will need to be developed and validated in a variety of settings.

AREAS FOR FUTURE RESEARCH

More research is required to build a new body of knowledge that will better bridge the theoretical concepts with the practical requirements of programme planners. At the very least, this research should focus on the following broad topics.

Elaboration of the patterns of population–pathogen interaction

Better field methods are required to characterise the sexual behaviours and network structures within a population. For example, there are few standard ways of summarising the nature and extent of female sex work within a population, although some approaches have been suggested. Similarly, rapid techniques to determine the size and nature of the...
interaction between core groups and bridge populations are lacking. Within the general population, more robust methods for describing the nature of longer duration partnerships are also required. These methods should address concepts such as concurrency, serial monogamy, and the time gap between relationships. This entails developing better survey methods and instruments for describing the distribution of sexual behaviours, especially among women. There is growing interest in the use of network analysis, but more widespread use of these techniques will depend on the development of better methods for classifying and interpreting these types of data.

Research is also required for a better understanding of the implications of various population–pathogen interactions on the epidemic trajectory. In this regard, mathematical modelling could be used to provide a richer understanding of the dynamic interplay between different sexual network structures and pathogens. This modelling could include simulations of various intervention strategies, implemented at different epidemic phases as a tool for understanding the impact of different phase specific strategies in various settings.

Empirical research is also needed. A systematic process of gathering and analysing empirical observations from diverse populations would provide insights into the epidemic patterns encountered in relation to the population characteristics. In addition, an elaboration of rapid epidemiological methods for describing the nature of the transmission dynamics for particular pathogens is warranted, particularly with respect to determining the extent to which transmission is dependent on core groups.

Development of rapid, accurate, and reliable methods for determining the epidemic phase

The various methods for determining the epidemic phase need to be better developed and tested in diverse epidemic circumstances. In this regard, it is important to know what can be inferred from static or trend data on the incidence and prevalence of a pathogen. Geographical methods for determining the epidemic phase also need to be elaborated. Comparative studies for different pathogens in diverse population circumstances could provide insights into the general application of geographical concentration as a proxy for epidemic phase. This will require the development of sound statistical methods for summarising geographical concentration. With respect to network analysis, efficient methods are needed to improve our understanding of the distribution of STI pathogens in spread and maintenance networks.

CONCLUSIONS

Few would argue that strategies for STI prevention and control should be appropriate to the phase of the epidemic. Yet the promise of theoretical advances in assisting in the application of phase specific strategies has been dampened by the lack of empirical data that can help to show how different epidemics will progress through epidemic phases, and how strategies should thus be planned. In this paper, some crude generalisations have been formed from some of the existing theoretical and empirical observations. Clearly, a great deal more refinement is required both in the classification of epidemic trajectories and in understanding the programmatic implications. The increasing interaction between theorists, mathematical modellers, and those engaged in observational research and programme planning holds great promise for this field. As new theoretical insights are gained, they should be aggressively pursued with empirical research so that the goal of attaining practical strategic approaches to STI prevention can be realised.

ACKNOWLEDGEMENTS

Many thanks to Sevgi Aral and Stephen Moses for their thoughtful reviews and suggestions on this manuscript. JB is supported by a career award from the Canadian Institutes of Health Research and the Manitoba Health Research Council.

REFERENCES