Anticipating Outbreaks
A Prevention Role for Integrated Information Systems

SINAN K. ARAL, MSc, BA,* STUART M. BERMAN, MD, ScM,† and SEVGI O. ARAL, PhD†

IN THE RECENT PAST, an outbreak of syphilis among middle and upper middle class teenagers living in Atlanta suburbs took most people, including public health officials, by surprise. This outbreak subsequently received extensive media coverage and left many parents, public health officials, and citizens to realize how little they knew about the sexual and social behavior patterns of their children. As is usually the case, only after the outbreak (which occurred in Rockdale County) was recognized were analyses done that identified sexual and social circumstances that permitted or resulted in the increased disease transmission. This episode may have implications for sexually transmitted diseases (STDs) and sexual/health behavior surveillance. Although identification of this syphilis outbreak was prompt, in general, broader syphilis epidemics have taken considerably longer to be identified and responded to. Integration of disease surveillance data with behavioral surveillance data, and their integration across local areas, may allow public health programs to be better prepared to document and respond promptly to emerging and reemerging infectious diseases.

The Rockdale County outbreak has been well analyzed. At the center of the outbreak was a group of young girls who, in various combinations, met periodically to use drugs and to have a variety of sexual interactions with several groups of slightly older boys. Investigations starting with a small number of syphilis cases revealed a highly interactive network of young individuals engaging in intense concurrent sexual activity.

We know more about the sexual and social networks that were involved in the Rockdale County outbreak than we do about most other outbreaks in the United States, even though the public health techniques employed in the investigation, i.e., cluster-interviewing, have been around for a long time. It should be noted that the significance of the demographics of the Rockdale County case-clustering was noted promptly by a nurse in a local STD clinic. However, such prompt recognition may well represent serendipity—discovery might not have occurred so readily had cases presented to several different venues. Especially since the United States has officially launched a National Plan to Eliminate Syphilis by 2005, syphilis surveillance in particular and STD surveillance in general should be reconsidered and systematically improved to facilitate outbreak detection.

An improved system of STD surveillance would take into account a number of recent developments in our understanding of the way people interact and the way behaviors and STD morbidity relate to each other. Recent changes in sexual behavior, rapid social and economic globalization, and fundamental developments in management information systems (MIS) could be incorporated into such an approach. Ideally, STD prevention programs would be able to use a wide variety of relevant information in planning and evaluating STD prevention strategies and in monitoring STD rates and identifying possible epidemics promptly. Some of this information may need to be collected by the STD programs solely for their purposes; additional information would come from other sources, collected for other purposes, but would be integrated with disease surveillance data for use by STD programs. To that end, we propose a modified approach to STD monitoring that combines behavioral, social contextual data with biomedical surveillance.

The authors thank Patricia Jackson for her outstanding support in preparing this manuscript.

Correspondence: Sevgi O. Aral, PhD, Associate Director for Science, Division of STD Prevention, CDC, 1600 Clifton Rd., NE, M/S-E02, Atlanta, GA 30333. Reprints: CDC, CHST, Information Technology Services, 1600 Clifton Rd., M/S E06, Atlanta, GA 30333.

Received for publication December 20, 2000, revised June 21, 2001, and accepted July 5, 2001.
Fig. 1. The STD surveillance matrix.

data, and integrates information systems across local, regional and national units.

The Content and Scope of STD Surveillance

It may be useful to consider surveillance of sexually transmitted infections (STIs) and STDs from two perspectives, namely content and scope. Surveillance systems should address at least four content areas: 1) incidence or prevalence of disease (e.g., AIDS, pelvic inflammatory disease [PID], syphilis by stage); 2) incidence or prevalence of infection (e.g., HIV, cervical chlamydia); 3) sexual risk behaviors (e.g., age at sexual debut, number of sex partners, demographic and risk discordance of partners, concurrency of partnerships); and 4) social context parameters (e.g., age composition, rate of unemployment, prevalence and kind of drug use, prevalence of commercial sex, sex ratio, prevalence of transients, housing/living conditions).

In terms of scope, surveillance systems can be considered to have four levels: local, regional, national, and global. The content and scope of surveillance jointly define the STD surveillance matrix (Fig. 1).

In the matrix, the data that are most consistently available reflect surveillance of the bacterial STIs (e.g., gonorrhea) and STDs (e.g., primary/secondary syphilis), at local (county to state), regional (groups of states) and national levels. As one moves from local surveillance towards global surveillance, and from surveillance of infections and diseases towards surveillance of social context parameters, examples of existing surveillance systems become increasingly scarce even though global surveillance of behavioral risk factors for other disease categories is not a novel idea. MONICA, ERICA, EURALIM, and SMI are examples of behavioral risk factor surveillance systems in the area of cardiovascular health that are at least regional and on the path to becoming global systems. Although surveillance of social context parameters has not received much attention in the literature, some work has described the correlation between the social environment and STI/STD rates, and increasingly the need for rapid community assessments and situational analyses is recognized and points to the existing information gaps in this area. There is a considerable amount of relevant data addressing social context, collected by a variety of agencies that is available for access and utilization by STD prevention programs. The most visionary aspect of the STD surveillance matrix, and the one that will be most challenging to construct, will be the global surveillance of social context parameters. Information accumulating over the past decade about global drug and sex worker traffic routes, spatial movement of businesses, businessmen, laborers, and sex workers, and the related diffusion of STIs across space point to the importance of monitoring social context parameters globally.

Content of STI/STD Surveillance Systems

Surveillance of STDs and STIs may be accomplished through case reporting, prevalence monitoring and special surveillance-related studies and surveys. Globally, most surveillance activity focuses on STI/STD surveillance, but does not differentiate between surveillance of clinical disease and surveillance of infection, which is typically asymptomatic. The differentiation between disease surveillance and infection surveillance is important because of the high frequency of unrecognized and asymptomatic prevalent and incident STIs. Since such asymptomatic infections account for much of the transmission and spread of STI, and
subsequent STD, it is important to monitor the prevalence and incidence of infections as well as disease, as reflected in the matrix.

Including behavioral surveillance in addition to morbidity surveillance may alert public health authorities to the emergence of behavior patterns capable of rapidly spreading STI, and to the emergence of initial cases of a particular STD. The need for behavioral surveillance is well understood among public health practitioners working in STI/STD prevention.15 However, behavioral surveillance should not be limited only to surveillance of risky sexual behaviors. Since STIs are sustained in populations by "core groups" defined not only by high risk sexual behaviors, but also by inadequate contact with the healthcare system,3,16 it is equally important to include health behaviors in STD behavioral surveillance.

In the United States information on sexual behavior is obtained through surveys of representative samples of the population at either the national17,19 or state level.20,22 These surveys are in general not integrated with concomitant STI/STD testing. With better integration of data, the cause-effect relationships between sexual behaviors of individuals and their STD acquisition could be assessed more directly.

The same sexual behavior may be associated with different levels of STI/STD risk in different societal contexts. Data from the National Survey of Family Growth (NSFG) indicate that black and white women with similar, moderate numbers (two to three) of sex partners face markedly different risks of acquiring pelvic PID, an STD, because of contact with different sexual and social networks carrying different prevalences of STI's.23 Similarly, data from the National Health and Social Life Survey (NHLS) highlight the importance of sexual mixing in explaining observed racial/ethnic differences in STI/STD rates in the United States.3,24,25

These considerations point to the need for the collection of behavioral and biomarker data in an integrated fashion. The acceptability of such collection has been established.26 Moreover, two national surveys, the National Health and Nutrition Examination Survey (NHANES) and the National Survey of Adolescent Males (NSAM), have collected data on both sexual behaviors and STIs.14 With the advent of computer-assisted self interviews to facilitate collection of behavioral data, and urine tests for sexually transmitted pathogens based on nucleic acid amplification methods to facilitate measurement of biomarkers, integrated behavioral-biomedical surveys/surveillance are increasingly feasible.

For STI/STD prevention programs to be better informed about factors relevant to the spread and control of STIs, public health officials should be aware of and able to monitor social contextual parameters. Factors such as age composition, gender composition, unemployment, poverty, and social and spatial mobility vary across localities and over time; such factors influence major determinants of STI spread, including duration of infectivity, transmission efficiency, and sexual contact rate.16 Timely information about these parameters may permit STI/STD prevention programs to better target interventions and more accurately anticipate changes in rates of STI spread. Such relevant data, addressing social context, are frequently collected by a variety of agencies and are available for access and utilization by STD prevention programs.

**Scope of STI/STD Surveillance Systems**

The horizontal axis of the STI/STD Surveillance Matrix represents the geographical scope of the surveillance system and movement along the horizontal axis requires an integrated expansion of surveillance efforts from the local to the global. As acknowledged, surveillance in the United States currently is conducted at the local, state, regional, and national levels. To be most useful, surveillance data should be integrated across the horizontal axis and, ideally, expands integrated disease surveillance beyond the local and national levels.

The uses for surveillance data differ across the local, regional, national, and global levels and data appropriate to the purposes at all levels should be available. At local and state levels, surveillance data are used to monitor disease trends, determine both priorities and the geographic and demographic targets of program effort (for outreach or screening efforts), provide a basis for decisions about resource allocation, identify potential and materialized outbreaks, interrupt the spread of infection via partner notification services, and assure that individual cases have received appropriate treatment. On the other hand, at the national level, data are used for large-scale trend analysis, for directing national strategy, for coordinating, as appropriate, local and regional efforts across municipal borders, and for informing an appropriate research agenda to serve STI/STD prevention.

**Decentralized Action Versus Integrated, Standardized Information: A Balancing Act**

Integration of STI/STD surveillance across dimensions of both content and scope is a daunting challenge. Nevertheless, the technical aspects can be facilitated by the use of modern MIS. However, creating efficient and comprehensive information systems for disease surveillance and prevention requires a balance of interests. Although standardization and integration of health information systems can increase efficiency and precision,27 balance is necessary; standardization could lead to rigid, centralized implementation of integrated information systems which can rob localities of the flexibility needed to respond quickly and efficiently to time-sensitive infectious disease threats. The
design of information systems for disease surveillance requires appreciation of divergent information needs and assuring appropriate flexibility and efficiency.

This tension between permitting decentralized action and imposing an integrated, standardized information system has been addressed in theoretical works by Walsham, Zuboff, and Borja and Castells, which examine the delicate balance between centralized and decentralized computing solutions for a variety of public administration information needs. Walsham points to the necessity of balancing interests not only in implementing, but also in designing and developing systems used by both a central authority and a local, action-oriented "periphery." The ability of decentralized systems to empower the periphery while maintaining efficient transmission of data required by the center, captures the essence of current information needs for disease surveillance in the United States and elsewhere. The development model suggested by Walsham advocates "a process of continuous learning" from local experiences in order to modify information transmission as needs change. The concept of continuous development with local input is directly applicable to disease surveillance efforts because the nature, transmission, and spread of diseases change rapidly and require flexible responses.

The objective is to build a dynamic, creative relationship between the local and the global, and to avert rigidity. The sharing of information in efficient ways provides critical support to disease control and prevention. Early warning of emerging and reemerging infections depends on the ability to identify the unusual as early as possible, and for this, information is essential.

The need for accurate, consistent data across all levels will require data integration, which must be considered and addressed from several perspectives—by content, by level, and across surveillance systems. Thus, ideally, STI/STD data would be integrated with data on risk behaviors and social context parameters (Fig. 2). As STI/STD surveillance systems evolve and improve, they will be characterized by "comprehensive integration"—high levels of integration in both scope and content. Creating such integrated data available on a global basis will be an enormous challenge. However, as various STI/STDs are eliminated from many local and national populations, the issue of reintroduction of infection will become a serious consideration, and will be a stimulus to the development and integration of global STI/STD information.

Certainly any effort to revamp and broadly upgrade information systems will require a substantial investment of resources. Public health budgets are precious and finite, and competition for resources drives political resource allocation decisions leading to budget battles within the context of power and authority in organizations. However, costly as it may be to implement the systematic changes discussed, the costs of not taking action are typically unappreciated. To address the benefits of efficient systems and the costs of maintaining inefficient systems, we further examine integration and standardization as possible mechanisms for efficiency gain.

Integration

The lack of systems integration increases training and manpower needs since public health staff must use and maintain several unrelated systems. In addition, data remain inconsistent and costly to extract and reproduce for the purposes of disease research or local disease control. For example, in the United States, surveillance is based on mandated reporting structures, under which reports flow
from the local level to the federal and then back to the local, often using data collection systems and software provided by the Centers for Disease Control and Prevention (CDC) or independently developed by state and local health authorities. However, these approaches have frequently been program-specific, resulting in a mass of disparate data systems that burden CDC’s state and local partners. There are at least five separate data systems for AIDS/HIV reporting alone and several others dedicated to the reporting of individual diseases or groups of conditions. Examples of currently operating systems which have not been integrated include: National Electronic Telecommunications System for Surveillance (NETSS), Sexually Transmitted Disease Management Information System (STD/MIS), HIV/AIDS Reporting System (HARS), Supplement to HIV/AIDS Surveillance (SHAS), Adult Spectrum of HIV Disease (ASD), Pediatric Spectrum of HIV Disease (PSD), Tuberculosis Information Management System (TIMS), and the Public Health Laboratory Information System (PHILIS).

During the last decade and particularly during the last five years, the public health community has raised concerns about the disintegrated surveillance systems. Currently, CDC has a number of projects in place, including the National Electronic Disease Surveillance System (NEDSS) and the Integration Project, which aim to create an infrastructure of connectivity across public health surveillance systems. Within the domain of STI/STD prevention programs, most existing systems were products of old technology; DOS-based products that needed to be upgraded to Windows-based systems (Scott Danos, personal communication, August 2000). For example, the STD/MIS is a DOS-based product. In addition, the information technology capacity in many STI/STD programs is very limited. Assessments of existing STI/STD surveillance systems indicate deficiencies in reliability, continuity, and accuracy, as well as capability.

A policy of systems integration would contribute to the efficient use of resources and create uniformity at the local and national level, making data exchange more productive. Integration addresses key issues of efficiency and flexibility, and highlights areas of local/national/global cooperation in the effort to save resources for effective priority setting.

**Standardization**

Integration of MIS across the content and scope of STI/STD surveillance necessitates the use of standard information data elements, flows, and processes. Currently, in the United States the lack of a standard procedure for laboratory and physician disease reporting, the existence of several “standard” codes (e.g., AMA/CPT codes and ICD9 codes for classification of diseases), the lack of unified behavioral data elements, case definitions, and network protocols make integration of national reporting systems extremely difficult. In addition to its effects on integration, the lack of standardization also contributes directly to inefficiency and disrupts the smooth flow of information. Standardization of information processes in public surveillance systems is critical to efficient information flows and comparability of aggregated data across regions, and would facilitate prompt identification and response, at all levels, to emerging disease threats.

Specific issues relevant to standardization include format, flow, and content. Format refers to the mechanics of data transfer, message coding, and standards of classification. Flow refers primarily to how information moves from user to user, including issues of authority and timeliness of data transfer. Content focuses on data elements and on exactly what types of information message transfers should contain.

To communicate efficiently, participants in an information exchange must, at the most basic level, speak the same language. Without a common vocabulary and standard coding structures and messaging formats, precise and effective communication is difficult. Private physicians, laboratories, public STD clinics, and public/private family planning clinics use different means and methods of reporting surveillance data to state and local health departments. Much of this incompatibility is influenced by the fragmented nature of the information systems used. In essence, prior agreement on ordered information exchange protocols are rare.

For example, although pilot projects have begun in the state of Washington to test the implementation of standard messaging, no national message standard exists to define messages for laboratory and other clinical results, immunization reporting, drug usages, patient registration, or clinical trials. Many advocate Health Level Seven (HL7) for implementation in all clinical data transmissions, and in addition to HL7, several standards exist for coding of data elements (the Logical Observations Identifiers Names and Codes [LOINC], the Systematized Nomenclature of Medicine [SNOMED], Current Procedural Terminology, version 4 [CPT4], and the National Library of Medicine’s Unified Medical Language [UMLS]). As a result of the diversity of coding and message standards, current disease surveillance and prevention systems are speaking dissimilar languages. The existence of so many divergent standards means that public health officials and other users either won’t have access to integrated data or that such data will only be obtained with difficulty and at substantial personnel and training expense.

Standardization of information flows is a necessary step in creating efficient, integrated information systems for disease surveillance. Public and private laboratories and physicians may need to report certain diseases and conditions to multiple jurisdictions. The data required to be reported is not necessarily consistent, in terms of content or format, and the method of transmission is also variable. There is varia-
tion among laboratories and physicians in levels of sophistication and method of reporting: some report electronically, while others rely on postal delivery, telephones, or fax machines. Electronic reporting was introduced into STI/STD surveillance systems in 1999, and now many, but not all, states report electronically (Scott Danos, personal communication, August 2000). It is believed that timeliness and completeness of reporting increases 30% when electronic reporting is introduced, however, electronic reporting raises costs by increasing the need for more sophisticated human resources. Even in those states that report electronically, physician-based reporting is poor and spotty. One surveillance goal is to improve physician based STI/STD reporting.

Not surprisingly, there is considerable variability across and within states in STI/STD reporting patterns. In addition reporting regulations vary substantially by state. For example, some states specify the laboratory tests to be used in the diagnosis of gonorrhea and others rely on clinical diagnosis; some states report diagnosed cases of chlamydia infection, while others report a case of chlamydia infection only if it is treated (Scott Danos, personal communication, August 2000). Furthermore, reporting regulations that exist are rarely enforced.

Both national assessment and local response would be facilitated by standardizing the reporting system and enhancing the efficiency of information flow. However, standardization should not be viewed as the ultimate solution. Attention must be directed to the "traditional aspects of power and authority in organizations" that Walsham and Zuboff describe.\(^{28,29}\) Clearly, an important obstacle to implementation of standardized information flow is the agreement that must be achieved among a daunting array of interested and powerful parties. However, cooperative and collaborative development of specifications between the national and local health authorities could facilitate implementation of principles agreed upon by all participants and allow for flexible evolution to new models of specification in the future.

Behavioral data elements also require standardization. While current information systems provide data to assess population-specific risks, integration of data between information systems is made difficult by a lack of comparability across data items that are intended to measure the same behavior (e.g., condom use). Standardization of content, addressing issues such as data elements, behavioral items and case definitions, would be a step toward having efficient, comparable data available for policy making.

Here again is a need to achieve compromise between conflicting needs and interests and available capacities. In this case, a balance must be struck between the need for comprehensive, uniform data and the variable capabilities of local health authorities, clinics, and physicians. Depending upon the condition under consideration (e.g., need for surveillance sensitivity versus specificity), minimum sets of data elements could be supplemented with larger and more complete sets on a case-by-case basis. The amount of demographic data required must also be considered in an attempt to minimize the burden of reporting while maintaining the prompt transfer of consistent and relevant information.

Conclusion

As public health information systems undergo transforming changes, STI/STD prevention programs must recognize how accurate, complete, and timely information can be utilized to achieve more effective STI/STD prevention; the programs must identify their information needs and advocate for and participate in the implementation of appropriate information systems. In fact, there is critical need for research to determine the extent to which better information would result in or permit improved STI/STD prevention or decreased morbidity rates. However, recent history would seem to suggest real utility; outbreaks across the United States of STI/STD (i.e., syphilis, gonorrhea, and chlamydia) among men who have sex with men revealed how surprisingly little we know about both morbidity and risk behaviors among this population.

The need for STI/STD surveillance systems that are integrated across scope (i.e., local to global) and content areas (e.g., risk behaviors and social context parameters) seems clear. Currently, a window of opportunity exists for creating the information systems that would best serve the purposes of STI/STD prevention programs. Changes will be made; those who are committed to improving the effectiveness of the nation's STI/STD programs need to assure that they are the right changes.

References