

# **Integrating Public Health Information and Surveillance Systems**

**A Report and Recommendations from**

**The CDC/ATSDR Steering Committee on Public Health  
Information and Surveillance System Development**

Spring 1995



CDC INFORMATION CENTER  
CENTERS FOR DISEASE CONTROL  
ATLANTA, GEORGIA 30333



## TABLE OF CONTENTS

EXECUTIVE SUMMARY .....	i
PREAMBLE .....	ii
INTRODUCTION .....	Page 1
I. INFORMATION REQUIRED FOR MONITORING PUBLIC HEALTH .....	Page 2
II. THE CHALLENGE: FRAGMENTATION, BURDEN, INFORMATION GAPS .....	Page 7
III. THE SOLUTION: INTEGRATED INFORMATION AND SURVEILLANCE SYSTEMS .....	Page 8
IV. CONFIDENTIALITY AND ACCESS .....	Page 16
APPENDIX A: Steering Committee Members .....	Page 22
APPENDIX B: Summary of Interim Work Group Recommendations .....	Page 25
APPENDIX C: Areas for Data Standards .....	Page 29
APPENDIX D: Public Health Information Issues Related to National Health Reform Proposals .....	Page 33
APPENDIX E: Definition of Terms .....	Page 42



# INTEGRATING PUBLIC HEALTH INFORMATION AND SURVEILLANCE SYSTEMS

## EXECUTIVE SUMMARY

The CDC/ATSDR Steering Committee on Public Health Information and Surveillance System Development was convened to begin implementing the highest priority objective of CDC's new strategic plan: the creation of integrated public health information and surveillance systems. This objective reflects the recognition that the current methods of data collection have placed a substantial burden on CDC/ATSDR's partners in state and local health departments. Given the increased emphasis being placed on improving customer service, it is important for CDC/ATSDR to look at current business practices and try to build upon existing resources in order to operate more efficiently.

Probably the most important factor influencing this Steering Committee has been the increased discussion surrounding health reform both at the state and federal level. Many of the proposals for health reform promote changes for the mechanisms in which health information is gathered. The Steering Committee recognizes that if data systems developed under any proposed health reform measure take into account the unique needs of public health, they will provide rich new sources of data for public health. However, these systems will still not be sufficient to meet all public health data needs. As a consequence, it will be necessary to establish and strengthen information systems to serve the unique public health data needs.

The report examines a) the information required to support essential public health functions and current public health practices of CDC/ATSDR; b) the challenges and solutions for integrating health information and surveillance systems; and c) issues relevant to confidentiality.

On the basis of this information, the Steering Committee developed a specific set of recommendations to guide the development, maintenance, and use of public health surveillance and health information systems at CDC/ATSDR. Three elements are essential to the evolution of integrated health information systems: data standards, a communications infrastructure, and policy-level agreements on data access, sharing, and burden reduction. Each element is necessary, but not sufficient, so steps must be taken to assure progress in all three areas. Although technical systems may be compatible, data will not be shared unless there are policy-level agreements to do so. Likewise, agreements on a policy level cannot be implemented unless adequate technical systems are in place.

Finally, confidentiality is critically important to public health, because many data necessary for public health practices are potentially sensitive. CDC/ATSDR will continue to use state-of-the-art methods to protect data and data systems and to maintain confidentiality of health information.

# INTEGRATING PUBLIC HEALTH INFORMATION AND SURVEILLANCE SYSTEMS

## PREAMBLE

The Centers for Disease Control and Prevention (CDC) is responsible for the preventing disease, injury, and disability throughout the nation. As the nation's prevention agency, CDC has led the fight against such diseases as malaria, polio, smallpox, toxic shock syndrome, Legionnaire's disease, and, more recently, acquired immunodeficiency syndrome (AIDS). CDC's responsibilities have expanded over the years and continue to grow as the agency maintains a full prevention and health promotion agenda. The agency focuses on prevention of infectious diseases, chronic diseases, injuries, workplace hazards, birth defects and disabilities, and environmental hazards. The agency also pursues improved quality of life for all by promoting healthy behavior and life-style choices and by fostering healthful environments.

As CDC's mission has evolved, the methods of accomplishing its mission have also become more complex. CDC works with many partners throughout the United States and the world. State and local health departments provide the infrastructure on which the health of Americans is built. Other partners include the individuals and institutions that educate and promote the health of Americans of all ages, such as school systems, local community groups, businesses, and voluntary and professional associations.

The science and practice of prevention have advanced greatly since CDC's beginnings in the 1940s, and the technology for diagnosis and treatment of diseases has become highly sophisticated. Likewise, the technology for processing and managing information has greatly altered the way CDC conducts business. In 1980, few staff members at CDC or in the states operated personal computers at their desks. Today, such computers, with greatly enhanced capabilities and with appropriate software, are not only common, they are one of the essential tools of public health practice.

CDC has a long-standing tradition of providing assistance to the public health community. As technology has evolved, CDC has provided new information management systems, in the form of dedicated data collection systems and software, to state and local health departments. The state and local health departments, eager to take advantage of new technology but often without resources to acquire and apply it independently, have adopted each new system. However, with time, the proliferation of new health programs and specific computer systems has led to a plethora of data systems, data bases, software, and reporting mechanisms; the burden represented by the resulting surveillance grid threatens to overwhelm CDC's state and local partners. In keeping with its role of facilitator and capacity builder, CDC must now maximize the efficiency of existing information systems while finding new, more innovative ways of conducting public health surveillance.

# INTEGRATING PUBLIC INFORMATION AND SURVEILLANCE SYSTEMS

## INTRODUCTION

The CDC/ATSDR Steering Committee on Public Health Information and Surveillance System Development was convened in the summer of 1993 to begin implementing the highest priority objective of CDC's new strategic plan: the creation of integrated public health surveillance and health information systems. Three issues have raised this objective to highest priority.

First, the mission of CDC is to promote health and quality of life by preventing and controlling disease, injury, and disability. In 1992, 13 percent of the gross national product in the United States was spent on health and medical care. Concerns about health care costs, disparities in access, and availability of health insurance, have fueled the ongoing national debate on health reform. Although discussions about health reform have focused principally on health insurance coverage, this debate has highlighted the dramatic impact that changes in the health care system are having on public health. Also, a critical challenge in public health is to ensure that those involved in government and market-led reforms have access to high quality data on which to base rational and effective prevention programs.

Second, the explosive development of information technology and its widespread and increasing use in the collection, management, and analysis of medical and public health data require that CDC/ATSDR change its data collection and data-management practices.

Third, this new information technology has allowed the creation of separate data collection systems by various CDC/ATSDR programs. CDC/ATSDR and other Federal agencies currently ask (and fund) their partner organizations to use specific software and information management systems. The variability in data collection systems, limited capacity for data management, and the burden of dealing with multiple software systems are major issues for these partners. Accordingly, representatives of state and local health departments continue to ask CDC/ATSDR to adopt a coordinated, streamlined approach to surveillance and health information systems. Finally, CDC/ATSDR has been asked to consider the widely varying and evolving data-management capabilities of its partners when devising data systems.

Given this high priority objective and the need to provide a solid foundation for building an integrated public health information system, in this report the committee reviews

- I. Information required to support essential public health functions and current practices of CDC/ATSDR;
- II. The challenges for integrating public health information and surveillance systems;
- III. The solutions for integrating public health information and surveillance systems; and
- IV. Issues relevant to confidentiality.

## I. INFORMATION REQUIRED FOR MONITORING PUBLIC HEALTH

CDC/ATSDR and its public health partners are concerned with a wide spectrum of health issues including infectious diseases, chronic conditions, reproductive outcomes, environmental health, occupationally related health events, and injuries. This array of problems requires a variety of intervention strategies for populations, in addition to the need to provide clinical preventive services for individuals. Some critical examples are the provision of prophylactic measures (e.g., vaccination, postexposure rabies prophylaxis), educational services (e.g., public health messages to diverse populations, counseling and prophylaxis for contacts of persons with certain infectious diseases), inspection of food establishments, and control of outbreaks. For these activities, the rational development of public health policy depends on public health information. For example, information on the age of children with vaccine-preventable diseases has been used to establish policy on appropriate ages for having vaccinations. Information on the prevalence of elevated lead in blood has been used as the justification for eliminating lead from gasoline and for documenting the effects of this intervention, and information on the rate at which breast cancer is detected has led to new policies regarding the recommended ages at which to have mammograms.

This variety of health problems and intervention methods requires different types of **information for action** and a broad array of data collection methods. Information for action must be useful to public health programs at local, state, and national levels.

### **Epidemiologic Foundation for Health Information**

The uses of public health data derive from an understanding of the basic epidemiologic parameters of time, place, and person.

- **Time:** Analysis of data over time can reveal trends in disease or injury. For all health conditions, a measurable delay occurs between the exposure and the problem. In the case of disease, an interval exists between exposure and expression of symptoms, as well as the interval between a) onset of symptoms and diagnosis of the problem and b) eventual reporting of the illness to public health authorities. For an infectious disease, this last interval may represent days or weeks, whereas for chronic disease it may be measured in years. Also associated with time is the ability to monitor trends and detect changes--a trend that shows an increase, which may require public health action, or a trend that shows a decrease, which may reflect the effectiveness of an intervention.

Another important term that is closely associated with the parameter of time is timeliness. Timeliness reflects the delay between any two steps in an information system. For some public health purposes, such as the detection of outbreaks, a rapid notification system is the necessary forerunner of an effective public health response. For example, a cluster of meningococcal meningitis cases among school children represents a public health emergency that requires immediate intervention. Other

public health actions may require detailed data but in a less urgent time frame.

- **Place:** The approach to the prevention and control of disease and injury is often determined by circumstances unique to the geographic distribution of the disease or of its causative exposures or risk-associated behavior. For example, elevated blood-lead levels in children may represent exposure to lead hazards in their environment and may require both medical and environmental intervention.
- **Person:** The characteristics of the people or groups who develop specific diseases or who sustain specific injuries are important in understanding the disease or injury, identifying those at high risk, and targeting intervention efforts. For example, disparities in health (incidence or severity of disease) among members of different population groups highlight the need to identify cultural, economic, or social factors associated with these health problems.

### **Types of Public Health Data**

For such activities, CDC/ATSDR and other public health agencies need at least seven categories of information: a) reports of health events affecting individuals; b) vital statistics on the entire population; c) information on the health status, risk behaviors, and experiences of populations; d) information on potential exposure to environmental agents; e) information on existing public health programs; f) information useful to public health but obtained by organizations not directly involved in public health practice; and g) information on the health care system and the impact of the health care system on health.

Reports of Health Events. Reports of cases of specific diseases of public health importance serve as the basis of many of CDC programs (e.g., the National Notifiable Disease Surveillance System, NNDSS). Public health information needs may dictate the level of detail needed in data collection. For example, on a weekly basis, the NNDSS seeks reports on all cases of over 40 conditions in the United States but collects only a small amount of information for each case, in order to minimize the burden placed on those who report. NNDSS data are used to monitor trends in disease, to evaluate public health programs, and to identify unusual occurrences of conditions that may require further epidemiologic investigation at the local level. For some public health purposes, however, effective action requires additional detail on each case.

For this reason, supplemental data collection systems have been developed for some of the diseases involved in the NNDSS. Such supplemental systems are sometimes less comprehensive in terms of the population represented but provide more detailed information on characteristics of the occurrence of disease. For example, cases of hepatitis are reported weekly to NNDSS for publication in the *Morbidity and Mortality Weekly Report (MMWR)*. In addition, the Viral Hepatitis Surveillance Project collects data on specific risk factors for different types of viral hepatitis in selected geographic areas. These data have been used to document the importance of behavior associated with sexual activity and drug use as a risk

factor for transmitting hepatitis type B and to target educational and vaccination programs. Other uses of data may require the ability to identify uniquely the patient whose case is reported and sometimes persons who have contact with the patient, as in the identification and treatment of persons in contact with cases of sexually transmitted disease (STD) or tuberculosis (TB).

State public health laboratories currently analyze 41 million specimens annually. Some of the data from these analyses immediately enter the electronic public health laboratory information system (PHLIS) and are used in monitoring both short and long range trends in the incidence of disease. In addition, private laboratories report several times as much data as the state public health laboratories do; most of this information is available in electronic form.

For diseases diagnosed through laboratory tests, data obtained from **laboratories** provide useful information about specific characteristics of a pathogen or toxic substance. For example, serotypes of *Salmonella* reported by laboratories can complement the use of data reported through NNDSS; such information is commonly used in identifying outbreaks that might otherwise not be detected.

Intervention and control of some conditions require more detailed information than can be obtained feasibly from a large group of clinicians or institutions. As a result, **networks of selected health care providers** have been organized to meet these targeted information needs. For example, CDC's Sentinel Event Notification System for Occupational Risks (SENSOR) targets select groups of health care providers as a component of a comprehensive approach that uses multiple data sources to provide information used in directing efforts to prevent workplace-related morbidity. The National Nosocomial Infections Surveillance System (NNISS) receives reports from a selected group of hospitals on the incidence and characteristics of hospital-acquired infections; data from this system have been instrumental in alerting health authorities to the emergence of antibiotic-resistant strains of bacteria, which in turn has led to the development of specific recommendations regarding use of antibiotics.

**Registries** provide detailed and periodically updated information on individuals. For example, ATSDR's National Exposure Registry is a system for collecting and maintaining information on persons with documented environmental exposure(s). The stated purpose of the registry is to aid in assessing long-term health consequences to the general population from exposure(s) to Superfund-related hazardous substances. This is accomplished through facilitating epidemiologic or health studies by a) verifying what are thought to be known adverse health outcomes (hypothesis testing), and b) identifying previously unknown, undetermined adverse health outcomes should they exist (hypothesis generating). This information assists ATSDR in providing advise on appropriate actions to be taken for a specific community.

Vital Statistics. The registration of all births and deaths is legally mandated in the United States. Because the systems established to collect these data also contain other health related information (e.g., birth weight, cause of death), they can be used in monitoring the public's health. For example, the mortality information system provides data (from death certificates)

on virtually all deaths and thus is extremely useful for assessing the impact of different causes of death and for establishing priorities.

Information on Health Status, Risk Factors, and Experiences of Populations. Since the determinants of many important public health problems are behavioral, health agencies need information that is not readily available from medical records on the prevalence of various types of behavior and on access to care. Thus, regularly conducted **surveys** of the general population are needed for public health. These surveys may range from large-scale assessments of the general population to assessments targeted at high-risk (i.e., particularly vulnerable) populations. This need is particularly acute at the state and local level. Surveys provide information on

- Baseline health status (e.g., the National Health and Nutrition Examination Survey, NHANES, and the National Health Interview Survey, NHIS)
- Morbidity (e.g., the National Ambulatory Medical Care Survey, NAMCS)
- Prevalence of specific behavioral risk factors (e.g., the Behavioral Risk Factor Surveillance System, BRFSS, and the Youth Risk Behavioral and Surveillance System, YRBSS) and medical risk factors (e.g., NHANES and Pregnancy Risk Assessment and Monitoring System, PRAMS)
- Use of health care services and identification of underserved populations (e.g., NHIS)
- Potential for exposure to toxic agents (e.g., the National Occupational Exposure Survey, NOES).

This information is used in developing prevention and control programs and in ensuring adequate delivery of health services.

Information on Potential Exposure to Environmental Agents. Information on exposures to environmental agents can be used in evaluating the risk to health represented by non-infectious diseases, injuries, and certain infectious diseases. For example, measurement of airborne particulates is useful in assessing risks related to certain pulmonary disorders (e.g., asthma and lung cancer). Information on vectors that may carry agents of infectious disease (e.g., ticks as vectors for Lyme disease and Rocky Mountain spotted fever, mosquitoes as vectors for viral encephalitides, and raccoons as vectors for rabies) is important in evaluating the risk of having such infection. Information on exposures to known risks supports the development and implementation of rational public health interventions (e.g., ATSDR's Hazardous Substances Emergency Events Surveillance [HSEES] System provides information on the public health consequences associated with the release of hazardous substances). In addition, information on exposures provides the basis for issuing alerts to the public and bulletins for clinicians on how to recognize and treat persons for health problems acquired through specific exposures.

Information on Program Management. Data necessary to operate public health programs include such items as number of clients served and cost of services rendered. These data are useful to public health officials in assessing the effectiveness of public health programs, comparing different programs, documenting the need for continuing a particular program, and maintaining accountability for tax dollars spent.

Information from Other Organizations. Data useful for public health are currently or potentially available from organizations whose functions may not be related to those of CDC/ATSDR and of state and local health departments. Data from the Bureau of the Census, for example, are necessary both for the reliable computation of rates and for the proper adjustment of rates for comparisons over time or in different geographic areas. The Environmental Protection Agency (EPA) compiles environmental air-monitoring data to assess compliance with standards for air pollutants established by the Clean Air Act. Data collected through this system are also used by public health officials for hazard alerts when pollutants exceed Federal standards and in studies of the effects of air pollutants on morbidity associated with respiratory diseases. The Occupational Safety and Health Administration (OSHA) and the Bureau of Labor Statistics compile data on the occurrence of work-related injuries and illnesses and exposure to hazards in the workplace which can be used for surveillance and research purposes. Similarly, many states compile workers' compensation claims data in administering their worker's compensation programs; these same data can be used for surveillance purposes. The Department of Transportation operates the Fatal Accident Reporting System (FARS), used in public health to assess risk factors for motor-vehicle-related injuries and deaths. The Federal Bureau of Investigation (FBI) crime statistics assist in evaluating the public health impact of intentional injuries, and the Consumer Product Safety Commission (CPSC) collects data on injuries related to consumer products.

Information on the Health Care System: Information is also needed on the health care system and the impact that changes in the system have on health. CDC provides a great deal of information to monitor the capacity of the personal health care system, utilization of that system, and access to health insurance and services by the American people. These data include: inventories of health care providers; surveys to determine patterns of utilization of health services, such as hospitalization rates and uptake of new technologies; tracking health insurance coverage on the part of the population and health insurance benefits provided by employers; and access to health care, and barriers (both financial and non-financial) to access.

## II. THE CHALLENGE: FRAGMENTATION, BURDEN, INFORMATION GAPS

Public health information exists in thousands of places: in the record systems of public health agencies and grantees, in the information systems of health care institutions, in individual case reports, and in the data files of surveys and surveillance systems. Many existing information systems were developed in response to high priority data needs, and they have continued to evolve in a manner that allows individual programs to deliver services and meet their information needs in a timely manner. These systems have been successful in enabling diverse programs to collect the information that they need to manage programs and provide client services, and direct prevention and control activities. The separation of systems has also fostered the protection of information security and confidentiality.

*However, because it is fragmented and compartmentalized, this information often cannot be aggregated to describe persons, populations, communities, or issues.* The development and evolution of these separate information/surveillance systems has resulted in a patchwork of data systems, which has in turn led to duplication of effort, left critical information gaps, strained cooperative working relationships, and made it difficult to accomplish the mission of public health.

Fragments of information on persons, communities, or topics are isolated in many different places. For example, a single patient may be treated by multiple providers, each with its own record system. Services provided for individual patients by public agencies may be recorded separately in the data systems of numerous categorical programs. Information needed to characterize the overall health of communities may be included in the records systems of health, environmental, social service, criminal justice, and other agencies, or may be collected through separate population surveys. This information often cannot be joined together because

- it has been collected in incompatible formats, using different definitions, personal identifiers, classification systems, or sampling strategies;
- the communications infrastructure is not in place through which data can be accessed, aggregated, and transferred; and
- policies, legislation, and organizational practices--although needed for the protection of confidentiality--may unnecessarily impede access to, and sharing of, information.

Since it is so difficult to join these data together, partners in the public health system cannot fulfill all of their information needs or efficiently manage the delivery of services. Public health agencies continue to create narrowly defined systems to collect the information they need for specific purposes, even though the information might be obtained more efficiently through other systems. The challenge now is to maintain responsiveness and protection of confidentiality and, at the same time, improve integration, efficiency, and usefulness.

### III. THE SOLUTION: INTEGRATED INFORMATION AND SURVEILLANCE SYSTEMS

*The joining of these fragments is the purpose of integrated public health information and surveillance systems.* Integrated systems can join fragments of information by combining or linking together the data systems that hold such information. What holds these systems together are uniform data standards, communications networks, and policy-level agreements regarding confidentiality, data access, sharing, and reduction of the burden of collecting data. CDC/ATSDR must take a leadership role in each of these areas to move public health toward integrated systems.

While the term "integrate" is used frequently in the field of information systems, it has many different meanings. The Steering Committee defines it as follows:

**Integrate:** to form into a more complete, coordinated entity by the addition or arrangement of diverse parts or elements; to incorporate these diverse elements into a larger framework, unit, or group to create a functioning whole and to eliminate redundancy.

The key element of integrated information systems is functionality rather than singularity. With integrated systems, a wide range of diverse individual information systems will continue to exist, but these systems must be coordinated, interconnected, comparable, and easy to use. Thus, the development of an integrated public health information system should

- build on data systems that have already been established, by modifying or augmenting existing surveys or surveillance systems;
- build on the patient care, client management, or other systems already in place or under development by health care providers, public health agencies, or others;
- build new systems when necessary to meet critical needs, but only within a larger, coordinated approach;
- minimize the collection and reporting burden placed on providers of data, as data should be collected once and then shared efficiently to meet all legitimate needs for these data;
- minimize the effort required to reformat, transmit, and share data with users, while protecting the privacy of individual data included in the system; and
- maximize the protection of confidentiality and minimize the potential for inadvertent, inappropriate, or other potentially harmful release of information.

A vision for an integrated public health information and surveillance system includes the following characteristics:

- A process exists through which users and providers of data develop a consensus regarding real needs for information, the best source (or sources, to allow for comparison) for each data element, and the development or modification of data systems, if necessary, based on agreed upon design principles;
- Information recorded by health practitioners, public agencies, laboratories, etc., as well as data collected by CDC/ATSDR (such as through surveys and records-based data systems) exists on a network that can be accessed by legitimate users in CDC/ATSDR, elsewhere in the public health community, and in the private sector;
- Information collected through different systems can be compared or linked to maximize analytic value and minimize the need for separate systems;
- Information is collected, maintained, disseminated, and used under strict guidelines for protection of security of data and confidentiality of individuals. As a general principal, information that could identify an individual would be maintained locally (e.g., in a state or local health department) rather than nationally. Strong sanctions against inappropriate release of information would be enforced; and
- Information that Federal grantees are required to report can be readily obtained from information systems that are designed to meet multiple needs (such as to support the delivery of services or to meet broad-based data needs) and do not constitute a separate data collection burden that competes for resources with the primary purpose of the grant.

Greater integration might be achieved through the consolidation of separate data systems, but this might not be the only means to the desired end. For example, it is difficult to imagine a single information system that would be comprehensive enough to meet all current needs and still be flexible enough to meet future needs. Integration can be achieved by having systems operate according to similar policies and standards so that they appear as one to the end user.

### **Recommendations Related to Integration**

CDC/ATSDR is one of many partners interested in the design of such an integrated system and is in a unique position to exert leadership in its development. CDC/ATSDR has specialized expertise in subject matter (disease etiology, prevention, health policy, etc.), in data collection methods (surveillance systems, surveys, etc.), and in the use of health information in public health practice and policy.

Other partners can help facilitate the development of integrated systems. For example, several organizations are pursuing the development of electronic patient records to support the

delivery of personal health services. Many public health agencies are also developing more comprehensive client-based information systems, and managed care organizations are seeking to integrate patient-care information from a variety of sources within their care networks.

Three elements are essential to the evolution of integrated health information systems: data standards, a communications infrastructure, and policy-level agreements on data access, sharing, and burden reduction. Each element is necessary, but not sufficient, so steps must be taken to assure progress on all three. Although technical systems may be compatible, data will not be shared unless there are policy-level agreements to do so. Likewise, agreements on a policy level cannot be implemented unless adequate technical systems are in place.

Develop and Adopt Standards. Integrated information systems require that users and providers agree on standards for factors common to many public health systems. This list includes common definitions of data elements and terms, common classification systems, compatible telecommunication protocols, and other technical specifications that allow different systems to be compared, linked, and otherwise integrated. The Steering Committee recommends that CDC/ATSDR should

- Take steps immediately to stimulate and facilitate the development of standards in the areas of a) core variables and other data elements, b) software applications, c) data transmission, d) data access, and e) confidentiality and security. In Appendix C, the Steering Committee has outlined specific sub-topics within these areas and recommends that a series of more specialized groups be established to develop standards for each area.
- Recognize the value of diverse approaches and skills within CDC/ATSDR and add value to public health information through the application of unique institutional and individual knowledge and expertise. To assure that diversity and innovation are not impeded, the Steering Committee recommends that CDC/ATSDR not adopt standards for the following areas:
  - selection of appropriate subject matter for surveillance or public health information, except to the extent that standards need to be developed for those elements common to many CDC/ATSDR systems. The Steering Committee feels that CDC/ATSDR programs possess subject matter expertise and should be charged with determining the legitimacy of their specific information needs within a general CDC/ATSDR policy framework.
  - methods, techniques, or approaches to data analysis. The Steering Committee feels that CDC/ATSDR programs are charged with the responsibility for assuring the scientific integrity of their analyses, and that it is not desirable to standardize specific analytic approaches.

- data processing, storage, and management. The Steering Committee feels that CDC/ATSDR has already established standards in these areas and that these standards are continuously updated by other standing committees within the agency. CDC/ATSDR should periodically revisit this issue to determine the appropriate time to develop standards in this area.
- data dissemination, including processes for review of manuscripts to be submitted for publication. The Steering Committee feels that each CDC/ATSDR program should be responsible for assuring that its products meet acceptable scientific standards.

Although CDC/ATSDR would not prescribe practices and standards in these areas, the Steering Committee assumes that individual CDC/ATSDR programs would adopt practices that are consistent with the goal of maximizing integration of data systems and sharing of data.

- Engage a contractor (possibly with 1% evaluation or other funding source) to expand the degree of detail included in the CDC/ATSDR Surveillance Coordination Group's "1994 Inventory of Public Health Surveillance and Health Information Systems." The contractor would
  - develop a data base with all information currently included in the inventory;
  - expand the inventory elements to include more detail on data sources, data elements, coding, etc.;
  - analyze the data base to determine the extent to which common sources serve multiple information needs within CDC/ATSDR; and
  - assist CDC/ATSDR in setting up an ongoing system for using the data base for reference, such as identifying information systems that might meet new needs or identifying opportunities for consolidation of similar systems.
- In setting priorities for standards development, focus particular attention on the standards that are needed to facilitate state-based efforts to develop integrated client-based information systems.

Provide for a Communications Infrastructure. An integrated public health information system is dependent upon an electronic information network infrastructure. Public health data tend to be distributed, i.e., they are created in numerous locations, for numerous reasons, and often remain as local data bases. Today many of these data are generated in electronic format. To facilitate rapid exchange of information, every public health worker in the United States should have the capacity to be linked to every other public health worker through telecommunications technology.

In its most general terms, an electronic communications infrastructure includes the equipment, protocols, and software that allow users to connect and exchange data with other users via

local- and wide-area networks. The means of communication are typically dial-up telephone connections for data communications but are moving towards local-area networks (LANs) that link the components of a group (such as a local health department) together. Wide-area networks (WANs) link these smaller components together with others involved in the same enterprise (e.g., linking local health departments with state and Federal public health agencies). Thus, sources of information become available to members of networks through their personal computers or other electronic equipment.

Numerous options exist to bring about local- and wide-area connectivity. Many organizations develop, own, and operate networks themselves, as is the case with CDC/ATSDR's local and wide area networks, CDC WONDER/PC and CDC INPHO. Privately managed network services, such as CompuServe and America Online, are examples of networks that provide both user interconnectivity and access to a mass of information. Other vendors now provide private WAN services by connecting corporate LANs. The Internet is a loose amalgam of thousands of computer networks that reaches millions of people all over the world. The users gain access to the Internet for a variety of purposes, including communicating with each other and obtaining a wide assortment of information and resources.

Regardless of the particular model for implementation, an essential component to any communications network is adherence to a set of protocols and standards that govern how connections are made between members of both local and wide area networks. The communications system should also be affordable and adaptable to the levels of need of its users. These principles need to be applied to CDC/ATSDR designed and managed systems (e.g., CDC WONDER/PC and INPHO) as well as to CDC/ATSDR's connections with other networks (e.g., the Internet).

To achieve linkage throughout the public health system, the Steering Committee recommends that CDC/ATSDR should

- provide leadership, resources, and technical assistance to the public health community to bring about universal interconnectivity.
- provide, through a standards setting process, communications protocols to be used for public health communications, including a comprehensive strategy for communications across the Internet (see also the discussion related to data and software standards, above).
- develop communications standards and promote their use by CDC/ATSDR software developers and by vendors that may develop software for public health applications.

Seek the Adoption of Policies, Legislative Mandates, Regulations, and Organizational Practices that Promote Data Access, Sharing, Burden Reduction, and the Protection of Confidentiality. An integrated public health information system is the result of agreements that exist between those who provide data and those who use these data.

Health information exists in a variety of settings: individuals have knowledge about their own situations, lifestyles, attitudes, etc.; health providers maintain patients' records about diagnosis and care delivered; laboratories report information on test results; government agencies routinely conduct surveys and generate information about their services. Users (e.g., public health officials, policy makers, researchers) describe a wide variety of needs for information (e.g., disease surveillance, epidemiologic and prevention research, trend statistics, policy research) and, to some extent, bear the burden of demonstrating that these needs are legitimate enough to justify the burden imposed on those who provide data.

In many cases, organizations serve multiple roles -- as initial sources of data, as intermediaries that add value before passing data on to other users, and as end users of data from a variety of sources. Currently, however, a variety of barriers exist that tend to divide and isolate users and sources, and have led to the creation of multiple, independent information systems to meet the needs of diverse users. Such barriers include policies, legislation, and organizational practices that unnecessarily impede access to data; end users that specify system requirements that exceed real needs, necessitating unique systems; categorical funding for surveillance and data systems; and a "turf and control" culture that encourages independence rather than cooperation.

An integrated public health information and surveillance system is based on a series of agreements that exist between those who have data and those who use these data. These agreements provide for the efficient flow of data to appropriate users by avoiding duplication, minimizing burden, protecting confidentiality, and maximizing analytic utility.

Barriers to this type of integration exist within CDC/ATSDR and between CDC/ATSDR and its partners. There is a need for CDC/ATSDR to establish a process to facilitate policy level integration among diverse CDC/ATSDR information systems and among organizations that CDC/ATSDR fosters or supports in the states. CDC/ATSDR must establish policies and an effective oversight mechanism to ensure that such barriers are removed.

A first step is for CDC/ATSDR programs to recognize the need to lessen these barriers and change the approach to designing and mandating new data systems. Within CDC/ATSDR, it is presumed that each program, by virtue of its subject matter expertise, should determine the legitimacy of its specific information needs. This determination should be made after a critical appraisal of the value of the information relative to the cost, the burden that would be imposed on sources of data, and the extent to which sources would be willing to agree to provide such data to CDC/ATSDR.

After a CDC/ATSDR program determines that information is required, the following steps should be taken in translating this need into the design of a new or modified information system:

- Are there other programs at CDC/ATSDR that should be involved in defining the information need and system?
- Are there others outside CDC/ATSDR, including states, that should be involved in defining the information need and system?
- Does similar information already exist elsewhere in an acceptable (or nearly acceptable) form?
- Can another system already in place be modified to accommodate the new need?
- Is the proposed source for the data already involved in providing data to CDC/ATSDR for other purposes, and could that mechanism (even for a different subject matter) be modified in a way that would be more efficient?
- Is the burden acceptable to providers of data?

The Steering Committee recommends the establishment of an **Integrated Health Information and Surveillance Systems Board** in order to provide an ongoing framework for CDC/ATSDR to exert leadership in the integration of public health information systems. The functions of the Board would be to

- promote the vision for integrated surveillance and health information systems and provide leadership for implementing this vision;
- provide institutional support for systems development and revision through such activities as
  - coordinating and overseeing workgroups for developing standards and for dealing with other topics
  - advocating for resources to support the development of CDC/ATSDR's integrated surveillance/health information system and advising CDC/ATSDR management on priority resource needs;
  - assuring that each CDC/ATSDR program with responsibilities for surveillance and health information systems establishes and maintains an internal process for adhering to standards and for reviewing exceptions to standards; and
  - promoting communication related to surveillance and health information systems throughout CDC/ATSDR.
- stimulate and sponsor innovation in surveillance and health information systems to ensure the adoption of cutting-edge methodology, including support for demonstration projects related to new sources of data or new methods of acquisition and analysis;

- assure customer (state/local agencies, other constituents) involvement in CDC/ATSDR decisions regarding surveillance and health information systems and attempt to identify and accommodate, to the extent possible, their key concerns and objectives;
- represent CDC/ATSDR surveillance and health information activities to relevant organizations and agencies in both the public and private sectors; and
- provide external evaluation of CDC/ATSDR surveillance and health information systems.

Membership to the Board would include

- senior representatives from all Centers, the Institute, the Program Offices (including CDC's National Immunization Program), and ATSDR;
- CDC's Office of the Director (Office of Program Planning and Evaluation, Associate Director for Science, Office of Minority Health, Office of Women's Health, and the Office of the General Counsel);
- two representatives from the Office of Program Support (one from the Information Resources Management Office and one representing the OPS Director);
- partner representation; a formal process would be developed to ensure that input from CDC/ATSDR's partners is received regularly. The function and role of these partners would be determined by the Board.
- outside constituents; the Board might also invite other individuals or organizations to provide technical assistance and participate in selected Board activities as deemed appropriate.

Existing surveillance and health information activities would be related to the Board in the following manner:

- The Surveillance Coordination Group could be replaced by the Board or could continue to function, but could become a subcommittee of the Board with newly defined responsibilities.
- The INPHO lead team should become a subcommittee of the Board.
- The INTERNET Steering Committee should complete initial strategic planning and transfer future steering functions to the Board.
- The Assessment Initiative workgroup should continue to function but should become a subcommittee of the Board.

- The Board would select individuals to serve as CDC/ATSDR representatives on various Department, PHS, or other workgroups and committees that address issues related to public health information and surveillance systems and/or data policy.

Additional steps should be taken within CDC/ATSDR, as well as in other public health agencies, that would include the identification of legislative provisions, regulations, administratively imposed grant requirements, and organizational practices that unnecessarily inhibit CDC/ATSDR's efforts to integrate data systems. These could be identified initially within each program, brought to the attention of Board; needed proposals for change could be developed. Similarly, CDC/ATSDR programs should examine the practices of grant recipients to determine whether CDC/ATSDR requirements are being implemented in a manner that restricts integration beyond requirements imposed by CDC/ATSDR.

Finally, CDC/ATSDR should support explicitly the development of integrated client-based information systems at the state level and should recognize important efforts by states to streamline both the collection of information and the delivery of public health services. Several steps should be taken in this regard: a) CDC/ATSDR should assure that its own requirements, regulations, and funding restrictions are not impediments to the development of such systems; b) CDC/ATSDR should encourage such development through direct funding (e.g., INPHO grants) and through creative approaches to using indirect and direct costs within existing grants; and c) CDC/ATSDR should facilitate the exchange of information on successful demonstration projects.

At the same time, it should be explicitly recognized that the states themselves have the lead in designing and implementing integrated client-based systems, and it should be acknowledged that it would be inappropriate for the federal government to mandate or to dictate the design of such systems.

#### **IV. CONFIDENTIALITY AND ACCESS**

Many of the data necessary for public health practice are potentially sensitive. For example, data relating to sexual orientation, potentially stigmatizing health conditions (e.g., human immunodeficiency virus [HIV] seropositivity), or illegal activity (e.g., illicit drug use) affect public health intervention efforts. To protect the privacy of individuals, data must be kept confidential. Yet, to protect the health of the community, health departments at the local and state level may have an urgent need for person specific information. The tension between the individual's right to privacy and the community's right to health creates a special issue associated with surveillance and health information. This, in particular, is an issue about which perceptions may be as important as reality in creating the balance needed for reaching decisions and setting policy. If the public believes that integrated information systems threaten confidentiality, such integration will be strongly opposed.

With appropriate safeguards for confidentiality, local public health officials need identifiable data to a) identify individuals who would benefit from treatment, prophylaxis, education, counseling and diagnostic testing for certain diseases; b) conduct special epidemiologic investigations; and c) conduct surveys.

Reporting of notifiable diseases provides an example of the use of identifiable information at the local or state level. States have the authority to mandate that physicians, laboratories, hospitals, and other health care providers report to local or state health departments the occurrence of certain diseases. At the local level, health departments often use this information as a basis for providing services to individuals. Reports of notifiable diseases may trigger epidemiologic and laboratory investigations to identify sources of disease, as well as the implementation of control and prevention measures. Knowing the identity of patients is essential to executing these public health functions.

With few exceptions, state health departments remove identifying information and assign case report numbers before they forward information to CDC/ATSDR. If, on review of case reports, CDC/ATSDR recognizes a case that needs further investigation, CDC/ATSDR contacts the state health department. The state then does the necessary follow-up.

In addition to notifiable disease reporting, health departments and CDC/ATSDR use various other methods to monitor diseases and injuries of public health importance. For example, disease or injury registries involve in-depth information collection, often including comprehensive follow-up information over the course of a case of disease. As mentioned previously, ATSDR's National Exposure Registry maintains a national registry of serious diseases and illnesses and persons exposed to toxic substances. This information assists ATSDR and state and local environmental representatives in assessing the long-term health consequences to the general population from exposure(s) to Superfund-related hazardous substances.

Outbreak or hazard investigations are conducted when there has been an apparent increase in the occurrence of a particular disease or injury, and a threat to public health is suspected. Occasionally, outbreaks or hazardous exposures investigated by CDC/ATSDR receive national attention (e.g., Legionnaires' disease, toxic shock syndrome associated with tampons, eosinophilia myalgia syndrome associated with amino acid supplements, *Escherichia coli* 0157 associated with hamburgers from a fast-food chain, hantavirus). But more often outbreak investigations are a less dramatic and more common part of routine health department activities. When CDC/ATSDR assists health departments with investigations, agency staff may act under the authority of local health officials whom they assist. CDC/ATSDR policy mandates that its staff retain individually identifying information only when absolutely necessary after leaving the health department setting.

CDC's National Center for Health Statistics (NCHS) conducts large-scale surveys to describe the health of the U.S. population. These surveys are based on carefully defined samples of individuals or health care providers and cover a wide range of health topics. These surveys

involve in-depth interviews and, in some instances, the review of medical records, collection of specimens, and testing of materials in the laboratory. In some surveys, individuals are re-interviewed to obtain additional or long-term data on health status. Thus, it is necessary to obtain and retain identifying information in order to recontact individuals for follow-up interviews or examinations. In addition, records may be subsequently linked with death certificates to assess the relationship between various health indices and survival.

In the record-based surveys, data on patients and residents are collected from facilities and providers (e.g., hospitals, nursing homes, physicians). These facilities and providers are guaranteed confidential treatment under section 308(d) of the Public Health Service Act for their own identification and that of the data they provide.

### **Uses of Non-Identifiable Data**

A variety of sources of non-identifiable data can be used by health agencies at the local, state, and national levels to monitor health. These include computerized vital records (e.g., birth certificates, death certificates) and medical care records (e.g., insurance claims, outpatient care records, hospital discharge records). Although this information may originally have been accompanied by identifying information, the identifying data are removed before they are disseminated. The non-identifiable data are then used to examine trends in causes of death, reasons for hospitalization, or the incidence or prevalence of various health problems.

### **Universal Identifier**

The National Committee on Vital and Health Statistics, an advisory group to the Secretary, Department of Health and Human Services (DHHS), has recommended the use of the Social Security Number (SSN) as the unique identifier for medical records. Currently, CDC/ATSDR uses the SSN in certain data collections and to link some data files. This number is obtained with the specific permission of the respondent. Respondents are advised that the data collected will be used for the stated purpose and that providing (or not providing) SSNs will have no effect on their social security benefits.

CDC/ATSDR is monitoring proposals for new unique identifiers. However, the confidential treatment of the SSN within CDC/ATSDR demonstrates that whatever unique identifier may be chosen in the future, CDC/ATSDR can and will maintain it in a confidential manner.

### **Protection of Data at CDC/ATSDR**

CDC/ATSDR maintains policies that govern the protection of person-identifiable and provider-identifiable data. These policies cover both the protection of paper records and the protection of computer-based records. CDC/ATSDR's mainframe computer centers at Research Triangle Park (NCHS) and Atlanta are guided by formal security manuals. Both centers have security manuals describing security policies and procedures.

All CDC/ATSDR electronic data and data systems are overseen by the CDC/ATSDR Information Systems Security Officer. The Security Officer chairs the permanent Subcommittee on Data and Information Systems Security of the Information Resource Management Coordinators Policy Committee. The focus of this subcommittee is managing the security of all CDC/ATSDR data computing facilities and networks.

Active discussions regarding technology to protect electronic records are ongoing. Groups such as the American Health Information Management Association continue to develop recommendations to safeguard the right to privacy and assure confidential handling of health information. These recommendations set confidentiality standards, policies for patient-access and data user responsibilities, education of users, formal confidentiality agreements, and physical controls.

Effective data security is provided by technical program safeguards and administrative and management policies and procedures regarding access. These safeguards include cryptography to encode data, personal-identification and user-verification techniques, access control software and audit trails, computer architecture designed to enhance security, and communications linkage safeguards:

- **Cryptographic** techniques encode data to provide confidentiality, to authenticate a message that ensures integrity, and to create digital signatures. **User verification systems** ensure that only authorized users gain access to the system. Access control techniques that rely on users' IDs also allow for restricting or mediating user access to specific data.
- **Audit trails** remain a standard form of security protection for computer systems. Robust security systems can provide a journal record of all accesses (authorized or not) to the system and can track to the specific data element.
- **Port protection** devices protect systems that require dial-up access. New security modems can combine passwords, dial-back, and/or encryption as a means of protecting network access to sensitive data.

### **Information Disclosure and Dissemination**

CDC/ATSDR disseminates data in reports and tabulations and through electronic means. These data are critical to health care researchers. CDC/ATSDR educates employees and reviews all products to ascertain that data that identify persons or establishments are not released inappropriately.

Employee Education and Responsibilities. CDC/ATSDR has prepared written manuals to describe the confidentiality guidelines applicable to its surveys and data collection efforts. All CDC/ATSDR employees, and contractors who handle confidential information are required, as a condition of employment, to sign a "Nondisclosure Statement." This statement discusses the

confidential nature of selected information covered by section 308(d) of the Public Health Service Act. It describes the penalties for unauthorized disclosure under Title 18, Section 1905 of the U.S. Code and the Privacy Act of 1974. In addition, the issue of confidential data is presented at orientation sessions for new employees.

Publication of Case Reports and Statistical Tables. The CDC publication entitled "Confidentiality" and "NCHS Staff Manual on Confidentiality" describe techniques to minimize the chance that information about an individual will be disclosed through the publication of statistical charts and tables. Case studies can only be published after all specific identifiers and demographic or geographic characteristics that could lead to identification have been removed.

Public-Use Electronic Data. CDC/ATSDR disseminates non-identifiable data through a public-use data policy. Often these data are from the national sample surveys and data collection efforts of CDC/ATSDR. All public-use electronic data are reviewed by the CDC/ATSDR Confidentiality Officer for adherence to the letter and spirit of Section 308(d) of the Public Health Service Act. Specialized non-public-use data are also released on a case-by-case basis and must meet the same requirements.

Authorized Disclosures. The Privacy Act authorizes the disclosure of identifiable information to the person to whom it pertains. Individuals or institutions may authorize disclosure by giving specific permission for the data to be released, usually for additional follow-up or research effort.

### **Recommendations Related to Confidentiality and Access**

- CDC/ATSDR programs use health information to monitor health trends, respond to urgent threats to public health, identify the causes of preventable disease and injuries, and develop and evaluate prevention strategies. The execution of these essential public health functions depends on CDC/ATSDR's continuing access to confidential health information. The agency recognizes the importance of individual privacy and data confidentiality and has successfully protected information in the past. CDC/ATSDR will continue to use state-of-the-art methods to protect data and data systems and to maintain the confidentiality of health information.
- Protecting the health of both the public and the individual is an important goal in the use of patient information. Accordingly, public health activities must be considered primary rather than secondary uses of clinical information. Information vital for public health should be collected in patient records.
- CDC/ATSDR supports Federal preemptive confidentiality legislation to provide uniform protection of records in all states and territories. This legislation must allow for disclosure of information, with appropriate protection, to public health officials for use in reports of disease outbreak investigations, public health surveillance, research,

and statistical analysis.

- Unique universal identifiers should be used on all health care records in order to link data and merge data files. The ability to do so is necessary for the protection of public health. CDC/ATSDR will maintain health information in a confidential manner, regardless of the identifier chosen.
- Health information collected by CDC/ATSDR does not generally include the personal identification of individuals. When collected as part of health monitoring or special investigations, such identifying information typically rests with CDC/ATSDR's partners in state or local health departments. However, an individual's identity may be inferred from a description of his or her demographic profile. For both identified and non-identified health information, CDC/ATSDR must continue to supply state-of-the-art security methods to protect the security of health information.
- CDC/ATSDR should actively influence discussions pertaining to protection of confidentiality and access to health data.

## APPENDIX A

### **CDC/ATSDR STEERING COMMITTEE FOR PUBLIC HEALTH INFORMATION AND SURVEILLANCE SYSTEM DEVELOPMENT**

Martha Katz (Chair)  
Associate Director for Policy, Planning and Evaluation

#### **Epidemiology Program Office**

Barbara R. Holloway  
Deputy Director

Edwin Kilbourne, M.D.  
Assistant Director for Science

Donna F. Stroup, Ph.D.  
Director, Division of Surveillance and Epidemiology

#### **International Health Program Office**

Melinda Moore, M.D.  
Deputy Director

#### **National Immunization Program**

Jose Cordero, M.D.  
Deputy Director

Donald L. Eddins  
Chief, Division of Data Management

#### **Public Health Practice Program Office**

Edward L. Baker, M.D.  
Director

David A. Ross  
Assistant Director for Information and Communications

#### **National Center for Chronic Disease Control and Prevention**

Gary C. Hogelin  
Assistant Director for Policy Planning, Office of Surveillance and Analysis

**National Center for Environmental Health**

R. Gibson Parrish, M.D.  
Chief, Surveillance and Programs Branch

**National Center for Health Statistics**

Jack Anderson  
Deputy Director

Edward Hunter  
Chief, Planning, Budget and Legislation Staff

**National Center for Infectious Diseases**

Ruth L. Berkelman, M.D.  
Deputy Director

Stephen M. Ostroff, M.D.  
Associate Director for Epidemiologic Science

Robert Pinner, M.D.  
Surveillance Epidemiologist

**National Center for Injury Prevention and Control**

Mark Scally  
Director (*Acting*), Office of Program Management and Operations

**National Center for Prevention Services**

Alan R. Hinman, M.D.  
Director

Samuel Dooley, M.D.  
Associate Director for Science (*Acting*)

**National Institute for Occupational Safety and Health**

Richard L. Ehrenberg, M.D.  
Senior Medical Epidemiologist

**Office of the Associate Director for HIV/AIDS**

James W. Buehler, M.D.  
Associate Director, Division of HIV/AIDS, NCID

**Office of Program Support**

Andrew Friede, M.D.  
Chief, Public Health Information Systems Branch, Information Resources Management Office

James D. Seligman  
Director, Information Resources Management Office

**Office of the Director**

Wilma G. Johnson  
Evaluation Officer, Office of Program Planning and Evaluation

Gaylon Morris  
Program Analyst, Office of Program Planning and Evaluation

Verla J. Neslund  
Deputy Legal Advisor to CDC

Dixie E. Snider, M.D.  
Associate Director for Science (*Acting*)

Emmanuel Taylor, Ph.D.  
Senior Epidemiologist, Office of Minority Health

**Agency for Toxic Substances and Disease Registry**

Jeffrey Lybarger, M.D.  
Director, Division of Health Studies

## APPENDIX B

### INTEGRATING PUBLIC HEALTH INFORMATION AND SURVEILLANCE SYSTEMS

#### Summary of Interim Work Group Recommendations

##### Recommendations Relating to Element 1 - Standards:

- 1) Develop proposed standards within CDC/ATSDR for items outlined at the end of this interim report and establish a process for expanding consensus on standards to include interested parties and partners. The following steps should be taken to begin the standards development process:

##### Establishment of Working Groups

- Circulate the list of proposed areas for standards included at the end of this interim report to the membership of the Steering Committee for agreement on areas for which standards should be developed. At the same time, ask Steering Committee members to indicate their interest in naming a representative of the relevant Centers, the Institute, or Program Offices (CIOs) to participate in each of the working groups.
- Determine whether there are any appropriate existing processes, groups, or CIOs that might be given lead CDC/ATSDR responsibility for a specific set of standards.
- Determine to whom workgroup recommendations should be made (e.g., the Steering Committee, newly established Board, Executive Staff, etc.)
- Name members to, and convene, the working groups.

##### Charge to Working Groups

The primary purpose of working groups for standards is to boost the standards development process by identifying existing activities, identifying potential partners, and doing the initial groundwork necessary for CDC/ATSDR to enter into a broader dialogue on developing consensus standards for public health. Specifically, the workgroups should

- identify other efforts at standardization that might be used as the basis for public health surveillance and information standards or that might constrain CDC/ATSDR's efforts to stimulate the development of standards.

- establish liaison with relevant organizations, agencies, etc., to obtain input into CDC/ATSDR's initial steps in the development of standards. Identify major stakeholders and attempt to learn (and accommodate, to the extent possible) their key concerns and objectives.
- develop a set of interim or proposed standards--in effect, CDC/ATSDR's proposals for standards--within 4 months after the workgroups are established. This timetable recognizes that it is important to put a reasonable set of standards on the table as soon as possible, but that there is the need for a wider consensus process to adopt a final set of standards that all of CDC/ATSDR's partners can fully implement. The interim set should maximize the prospects for eventual consensus by considering the views of CDC/ATSDR's partners; however, working groups are not charged to achieve consensus on a final set of standards within the time allowed.
- make recommendations for a process for gaining a wider consensus of interested parties and stakeholders and for a process for updating of standards as the environment changes.

There are separate forums within CDC/ATSDR for considering data-related activities, including the Excellence in Science Committee (integrity of CDC/ATSDR data and Internet policy), the group that is working on a CDC/ATSDR report on confidentiality, and oversight processes associated with INPHO. CDC/ATSDR needs to make sure that it coordinates its processes for coordinating data.

- 2) Do **not** develop CDC/ATSDR-wide standards in the area of appropriateness of subject matter. Further, permit CIOs to adopt their own practices in the areas of analytic methods; data processing, storage, and management; and publications peer review, consistent with accepted scientific standards and practices. The working group assumes that each CIO would, to the extent possible, develop such practices in a manner that maximizes the integration of data systems and the sharing of data (e.g., consistent with the common file formats and documentation standards discussed at the end of this interim report).
- 3) Engage a contractor (possibly with 1% evaluation or other funding sources) to expand the degree of detail included in the Surveillance Coordinating Group's "1994 Inventory of Public Health Surveillance and Health Information Systems." The contractor would
  - develop a data base with all information currently included in the inventory;
  - expand the inventory elements to include more detail on data sources, data elements, coding, etc.;
  - analyze the data base to determine the extent to which common sources serve multiple information needs within CDC/ATSDR; and

- assist CDC/ATSDR in setting up an ongoing system for using the data base for reference, such as identifying information systems that might meet new needs or identifying opportunities for consolidation of similar systems.
- 4) In setting priorities for the development of standards, CDC/ATSDR should focus particular attention on the standards needed to facilitate state efforts to develop integrated client-based information systems.

#### **Recommendations Relating to Element 2 - Communications Infrastructure:**

- 1) CDC/ATSDR should provide leadership, resources, and technical assistance to the public health community to bring about universal interconnectivity.
- 2) CDC/ATSDR should provide, through a standards-setting process, communications protocols to be used for public health communications, including a comprehensive strategy for communications across the Internet (see also Element 1).
- 3) CDC/ATSDR should develop and promote the use of communications standards for use by CDC/ATSDR software developers and by vendors that may develop software for public health applications.

#### **Recommendations Relating to Element 3 - Data Access, Sharing, and Burden Reduction:**

- 1) Come to closure on the nature, membership, and authorities of the process proposed for coordinating surveillance within CDC/ATSDR. Although the working group did not have time to fully debate the details of the process, there was a general consensus that this process should be generally representative of the CIO's, should be adequately supported by full-time staff, and should take an active role in representing to other agencies CDC/ATSDR's approach to integrating information systems.
- 2) Clarify the responsibilities and coordination of the various groups currently involved in developing CDC/ATSDR policy on information systems (i.e., Excellence in Science, Steering Committee, INPHO oversight, new Board, etc.)
- 3) In translating the **need** for information into the design of a new or modified information **system**, CIOs should follow a series of steps designed to minimize duplication and maximize prospects for integrated systems. These steps are outlined in the body of the report.
- 4) Identify legislative provisions, regulations, administratively imposed grant requirements, and organizational practices that unnecessarily inhibit CDC/ATSDR's efforts to integrate data systems and develop proposals for change.

- 5) CDC/ATSDR should explicitly support the development of integrated client-based information systems at the state level and recognize important efforts by states to streamline both the collection of information and the delivery of public health services. CDC/ATSDR should take several steps in this regard: a) CDC/ATSDR should assure that its imposed requirements, regulations, and funding restrictions are not impediments to the development of such systems; b) CDC/ATSDR should encourage such development through direct funding (e.g., INPHO grants), creative approaches to using indirect and direct costs within existing grants; and c) CDC/ATSDR should facilitate the exchange of information on successful demonstration projects.

## APPENDIX C

### AREAS FOR DATA STANDARDS

#### Common Data Elements/Standard Core Variables

- CDC/ATSDR should identify all of the forms, questionnaires, software, data bases, etc., used in its data collection activities (e.g., public health case reports, surveys, vital records system, program management data systems)
- CDC/ATSDR should then identify common and unique data items collected through these activities.
- For each data item that is collected in two or more data collection activities, CDC/ATSDR should compare the formats and classification and coding schemes used by each activity.
- Using the information from the last step as a starting point, CDC/ATSDR should develop a model format and classification scheme that could be used by different activities so as to allow collection and linking of data without loss of information or detail.
- To the extent possible, these model formats and schemes should comply with existing standards developed by the federal government and recognized standard-setting institutions.
- Model data-item formats and classification and coding schemes should be flexible and robust enough to accommodate emerging public health problems and evolving computer and data-management technology.

#### Software Development

Wherever possible, commercially developed software products should be used; CDC/ATSDR should develop software only as a last resort when commercially designed products are cost-prohibitive or fail to meet the public health need. When developed by CDC/ATSDR, software should have

- minimal specifications (e.g., systems should specify file format only);
- consistent interfaces (e.g., look and feel);
- standard file formats, operating system requirements, and minimal hardware requirements and documentation;
- standards for and availability of technical support.

## Transmission

One important aspect of integrated health information systems is the entry of data into systems, dissemination of data between systems (e.g., shared data base or linked data bases), and dissemination of information from systems. These connections are dependent on transmission method, transmission protocols, and standardization of data elements.

- Transmission method is the avenue by which data are conveyed. There are two basic transmission methods: telephone dial-up and networks. CDC/ATSDR needs to provide both methods for the near future until the public health community and the reformed health care system are fully connected to the information superhighway. There are two types of the dial-up method: computer-to-computer (e.g., bulletin boards or dial-up e-mail systems) and human-to-computer (e.g., telephone interactive voice response and interactive facsimile response). With the networks method, computers can exchange data across a connection that is constantly available, i.e., the physical connection does not have to be "established" through dial-up means upon need.
- Transmission protocol is the means by which a network handles its electronic traffic, including how messages are packaged, unpackaged, and interpreted. With the dial-up method, there are various dial-up protocols for modem, facsimile, and analogue voice transmissions (e.g., HST, CCITT Group IV, Touchtone). Likewise, networking protocols allow for the computer-to-computer transmission and interpretation of data without human intervention (e.g., IPX, TCP/IP).
- Standardization of data elements is critical in integrated transmission and communications. If the sending and receiving sites do not share the same data element standards (i.e., form, format, sequencing, etc.), extensive manual intervention is required to route, interpret, and merge the data into the system. Examples of data standardization for transmission include ANSI X.400 and X.500 for electronic mail, SMTP for Internet mail, and a host of X.12 formats for Electronic Data Interchange (EDI) of commercial business data.

Once the elements for integrated transmission of data are met (i.e., use of standard transmission methods, protocols, and data elements), the final step in full communications integration is linkage of communications pathways. Integration of communications can certainly be achieved by providing singularity of pathway; however, this is not necessary to achieve a useful level of integration from the perspective of the user. For example, on the Internet, information servers all around the world appear integrated to the user because they are linked together into a World-Wide Web through the Mosaic interface. Currently, most of CDC/ATSDR's communications systems are stand-alone and do not provide this level of integration. Some exceptions include the National Electronic Telecommunications System for Surveillance which facilitates integration of multiple programmatic surveillance data sets to CDC; the CDC Voice/Fax Information System (CDC VIS) which integrates the voice and fax

access to CDC/ATSDR public health information; and CDC WONDER, which also provides e-mail capability and access to CDC information.

### **Data Access**

In order to promote the sharing and linkage of data and to maximize use of available data, CDC/ATSDR should adopt a series of practices and standards regarding formatting and accessibility of data. Areas to be specified include

- for each information system, there should be a public use release of data at the most detailed level (subject to confidentiality requirements, use agreements with providers, etc.).
- informed consent statements and agreements used to collect data should not restrict appropriate access.
- there should be standards for timeliness of release and editing standards, quality levels, etc.
- there can be different standards for access within CDC/ATSDR as well as standards for access outside CDC/ATSDR.
- the most detailed level collected for each data element should be preserved and made available to users (e.g., single year of age should be available on a data tape, rather than recoded age groups)
- each data system should follow common documentation standards, including any federal standards developed through the National Information Infrastructure (e.g., collection, coding, imputation, file formats).
- an inventory of available public-use data should be developed and disseminated.

### **Confidentiality/Security**

- Development of standards should be coordinated by appropriate CDC/ATSDR staff (Privacy Act Coordinator, ADP Systems Security Coordinator, the Office of General Counsel, the Human Subjects Review Coordinator, etc.) and existing committees.
- Reinforce CDC/ATSDR policy **not** to collect personal identifiers unless absolutely necessary. Develop criteria for the collection of personal identifiers.
- Develop algorithms for the creation of identifiers when those other than, e.g., name or SSN are needed.

- Develop common methodology for encryption for each transmission (e.g., RSA public-key encryption).
- Develop procedures for the publication of data that assure that identity will not be disclosed inadvertently.

## APPENDIX D

### PUBLIC HEALTH INFORMATION ISSUES RELATED TO NATIONAL HEALTH REFORM PROPOSALS

November 1993

**NOTE:** This document was drafted by a working group of the Steering Committee in November 1993 to consider the implications of national health reform proposals, (e.g., President Clinton's Health Security Act), to explore potential public health applications for information systems included in those proposals, and to recommend potential steps that could be taken to influence the direction of discussions of such systems. While broad-based reform proposals are no longer being actively debated at the national level, the issues outlined in this report may have application to future debate at the national level as well as to current considerations in states.

#### **Proposed Components of a National Health Information System in the Health Security Act**

Collection, management, analysis, and use of the enormous quantity of information needed for managed-care delivery would require new data systems. Accordingly, the proposed Health Security Act devotes considerable attention to the development of a national health information system. Only a general outline of the functionalities of this system is described in the President's proposed legislation. However, the proposal specifies that a National Health Board would oversee establishment of the network's precise specifications.

The network would have the three components described below.

##### **Enrollment Data**

Health Alliances would maintain and update each year information on each person eligible for health insurance coverage. At a minimum, this information would include identifying information (e.g., name, address, race, ethnicity), source of coverage, and any information required to adjust premium payments to health plans based on the individual's risk (e.g., age, sex). The Alliance would issue to each enrollee a Health Security Card with a unique identifying number.

The utility of the enrollment data base for surveillance and other public health functions depends on whether the enrollment form includes information beyond this minimum level, such as demographic information (e.g., race, ethnicity, occupation) or additional information on personal risk (e.g., behavioral risks, medical history).

### **Encounter Data**

A standard, minimum set of data similar to that now provided on claims forms would be collected for each encounter with the health care system. Each Health Plan would be responsible for collecting these data from each affiliated provider and for providing data to the network. An example of the level of detail that might be mandated for physician contact is the current HCFA-1500 form, which provides for itemization of diagnosis, procedures, and disposition. Hospitals would report data similar to those on discharge abstracts.

The utility of the encounter data system to public health depends on the accuracy of data, the length and content of the minimum data set, the extent to which an encounter might be used to obtain risk factor or other information not usually recorded during the clinical encounter, and issues relating to privacy/access.

### **Administrative Data**

To facilitate certification and other functions of Health Alliances and public agencies, data would be collected on characteristics and operations of health plans and providers. Some of this information (e.g., staff ratios, percentage of physicians who are board certified, percentage of enrollees who leave the plan each year) may be provided to consumers through the Quality Management Program to aid them in choosing among competing health plans.

The Health Security Act proposes that data from each of these systems be aggregated and processed in regional data centers supported by the Federal government. These centers would then provide access to data to the Plans themselves, to Health Alliances, to state and Federal agencies, and to researchers in accordance with privacy legislation to be proposed by the National Health Board.

In addition, the Health Security Act also calls for surveys of individuals to measure factors that would not be available through data in the network. These consumer surveys would obtain information on consumers' satisfaction with Health Plans, access to care, and other items to be used in choosing among competing Health Plans. Large samples would be involved (potentially 1 million interviews annually), and the potential for small-area (geographic) estimates makes this data source important for meeting public health information needs that are not being met with existing data systems.

Health reform would entail substantial changes and revisions in the current health care system. However, at this time it is difficult to predict what exact effects the reform efforts would have. Nevertheless, current issues surrounding public health information and surveillance systems at CDC/ATSDR must be addressed in the context of information likely to become available because of health reform. Health reform, either as contemplated in the President's proposed Health Security Act or as outlined in other proposals, would likely lead to centralized or regional reporting of enormous amounts of health-related and administrative information. Since plans for health reform data systems are currently being developed, an

even more pressing issue is how CDC/ATSDR and other public health agencies can participate in designing the new data systems and how to ensure that the needs of public health for data and information are given proper consideration. Thus, as currently contemplated, a national health information system called the Network in this report) would be created under health reform that would be broad and comprehensive.

Data collected at the time an individual enrolls in a Health Plan have the potential to form a virtually complete registry of U.S. residents (lacking only information on persons not covered--e.g., undocumented aliens). Nevertheless, it is impractical to think that anything more than extremely abbreviated clinical data on each encounter with the health care system could be compiled at the national level. In effect, at the Federal level, the network would yield data that are broad in scope but not necessarily comprehensive enough for all public health uses--in effect a system that can be described as being a "mile wide and an inch deep."

This "mile" of data can fulfill many public health information needs in two principal ways: a) by providing a broad picture of the population, including reasonably complete coverage of diagnoses and the delivery of services of public health interest, and b) by identifying certain health events of public health importance (e.g., notifiable diseases) and then serving as a pointer to detailed medical records for further investigation of these events, with appropriate safeguards for confidentiality. Nevertheless, the information needs of public health would not be totally met by the currently proposed health reform data systems.

## **Opportunities Associated with Public Health Data under Health Reform**

The medical-care system is primarily focused on treating persons for medical problems. The public health system emphasizes the prevention of disease and injury, the promotion of health, and the protection of the environment and of personal health. Although the medical care system provides medical services to individuals who seek care, activities of the public health system extend beyond the boundaries of individual providers and facilities to encompass the entire population--hence the term "population-based services." Health and quality of life for populations would become realities only with interaction and coordination throughout the medical care and public health systems. In a restructuring of the health care system, public health must be able to continue to provide population-based services and take advantage of new opportunities.

### **Reports of Health Events**

An automated health reform data system could address the underreporting of many diseases by generating a report of any instance of a notifiable disease. Reporting could be expanded easily to include diseases or conditions for which reporting is not currently practical (e.g., upper respiratory infection, traumatic brain injury). Linkage of the health reform network to more detailed data systems (e.g., records of patients) would augment information available through the network.

The linkage between outcomes and enrollment data (e.g., age, sex, and geographic location) would provide the information needed for a virtual **registry** for any disease or health outcome that requires relatively frequent contact with the health care system. For example, for diabetes mellitus, the network would initially detect both incident (new) and prevalent (previously diagnosed) cases. However, after 1-3 years of data history, patients who receive a diagnosis of diabetes would probably represent only new cases. This diagnostic information--together with data on age, sex, and geographic location--would permit the calculation of incidence. The current debate over the utility and cost-effectiveness of a diabetes registry would become moot; the data would be available when needed.

### **Vital Statistics**

The broad population coverage of health-reform plans, coupled with information from encounter data, provides the opportunity to obtain some of the information currently collected through the vital records registration system. Although the current registration system would likely continue to be used to account for events occurring outside the health care system or among persons not covered by health insurance, some medically related information might be obtained more efficiently as a by-product of the health information system.

### **Information on Health Status, Risk Factors, and Experiences of Populations**

Since population registries are rare in the United States, current health surveys rely on costly or complex sampling approaches. A national health information system provides the opportunity for population-based sampling from the enrollment and encounter data base. If enrollment data bases are to provide useful information, they must be updated periodically. Storage of data should permit flexibility, allowing for sample selection from specific strata or subsets of the population (e.g., sampling of individuals by age, race, sex, geographic location, economic status [as indicated by subsidized coverage]). In addition, samples of administrative records can be used without costly and time-consuming construction of sampling frames. Finally, it would be possible to sample on the basis of conditions, diagnoses, or procedures to obtain more detailed information on specific conditions and to conduct research on outcomes.

The data system would also facilitate long-term studies, in which individuals are enrolled in cohorts according to characteristics noted in the enrollment files (e.g., occupation). Follow up of these cohorts can then be conducted through surveys and analysis of encounter data. Further, an efficient linkage of health, risk, and behavior information from surveys to data on utilization and outcome in encounter files would enable the ultimate health effects of behavioral risk factors to be analyzed.

### **Information on Program Data**

The encounter data system provides the opportunity to monitor the use of clinical preventive services, their effectiveness, their cost-effectiveness, and their clinical and economic benefits. Such information can be used in the development of revised guidelines on practice.

## **Limitations of the Health-Reform Data Network for Public Health**

Although the health data system provides the opportunity for substantial gains, any system based on individuals and their encounters with the health care system has inherent limitations for public health purposes. In this discussion, general concerns for any type of public health data are discussed and data gaps for public health under health reform as currently contemplated are listed.

### **Data Quality**

Experience with nonfinancial uses of claims data gives rise to concerns about misclassification, incorrect coding, underreporting, and completeness of data. These problems arise, in part, because there are financial incentives to inflate severity. Without concerted efforts to validate data and to reinforce the importance of accuracy and completeness (e.g., training personnel and placing negative consequences on incorrectly reported data), diagnosis and procedure data on encounter networks may be difficult or impossible to interpret.

### **Timeliness**

The volume and complexity of records that would be entered into the system would pose challenges in processing and reducing data to a form usable for analysis. For data that are time-sensitive (e.g., identification of cases of a notifiable disease that require immediate intervention), recognition of the event at the local level is critical. It may not be possible to obtain such information reliably from central processing centers in the requisite time frame.

### **Detail**

Compromises required to achieve an efficient and acceptable minimum data set would inevitably lead to a loss of detail. For example, information on risk factors that is necessary for the intervention and prevention of conditions may not be readily available.

## **Essential Public Health Data Unavailable from the Health Reform Information Network.**

Information beyond that available from specific encounters of patients with the health care system is essential for public health practice, including

- community characteristics (e.g., climate, rural/urban location, occupational exposures, air pollution levels) and interventions (e.g., seat-belt and helmet laws) that may affect health but would not be documented in a clinical setting.
- unique characteristics or circumstances of special population groups (e.g., non-traditional providers, cultural factors in seeking care) that are not feasible to include in standardized data sets designed for the overall population.

- social predictors of public health problems (e.g., poverty).
- knowledge, attitudes, and practices of the population toward health, public health, and health care.
- risk-associated behavior of the population that may not result in an encounter with the health care system until years in the future or behavior that is poorly documented in the resulting medical encounter (e.g., high-fat diet).
- events of public health importance that happen to individuals but do not result in health care encounters (e.g., fatal injury that occurs outside the hospital).
- conditions, illnesses, or risk factors that are undiagnosed or unattended by a health practitioner (e.g., seroprevalence surveys).
- quality of air, water, food, or blood supply.
- data on geographic distributions of disease-vector populations.
- information on highway safety.
- health conditions that do not require encounters but are of concern because of the morbidity they cause or their economic impact on the population (e.g., upper respiratory infections, minor injuries).
- persons who have encounters, but choose to pay using their own funds or to use free services even though they have insurance coverage for the services (e.g., anonymous human immunodeficiency virus (HIV) testing, free clinics, out-of-pocket or out-of-plan use).
- persons who have health conditions that normally would require encounters but who have barriers to access.
- persons who are not covered by the system (e.g., undocumented aliens).

Beyond data on individuals, populations, and their environment, the following information on health care services and delivery is also critical to public health:

- data on the ways in which providers organize practices and deliver care.
- information on the extent and nature of primary-care services that continue to be delivered by public health departments.
- information supporting certain specific aspects of the practice of public health (e.g., case finding during outbreaks, tracing contacts of persons with STDs, household intervention in situations involving lead intoxication).
- services of public health importance delivered by providers but not covered by health plans (e.g., substance abuse treatment not in basic benefit package).
- nature and extent of state and local population-based programs (e.g., smoking cessation, health education, or health screening).
- information on the legal and regulatory environment in which health care and public health are delivered.

## **Recommendations Related to Health Reform**

After determining its specific information needs, CDC/ATSDR should decide which information would likely come from health reform data systems and which information would be needed from other systems. Data collected through health reform systems are likely to be limited to a small amount of data from all enrollments and encounters. At the same time, access to more complete files must be granted to ensure that CDC and other public health organizations can obtain greater detail when needed. Finally, since health-reform data systems would not be able to provide part of the data needed for public health, CDC/ATSDR and other public health agencies must continue to develop and maintain other sources of data.

### **New Partners**

CDC/ATSDR and its existing collaborators in state and local health departments should establish new partners under health reform (e.g., health insurance organizations, health alliances, health boards, consumer groups, and private medical providers). Specifically, jointly with state agencies, CDC/ATSDR should conduct demonstration projects to assess the utility of public health information and surveillance systems linked to health care-delivery systems and should provide a model for a public health data system after health reform. Potential project sites include states which have enacted health reform legislation and large managed care systems or health maintenance organizations. Sites should be chosen so that existing funding can facilitate early completion of the project (e.g., Robert Wood Johnson Information for State Health Policy project sites or Assessment Initiative states).

### **Overall Requirements for the System**

CDC should advocate that a health reform data system must provide for

- timely processing and access, including special handling and real time access to data on sentinel events and access by public health agencies to the data base for sampling, identification of individual cases to facilitate public health follow-up, and guidelines for disclosure that are responsive to state level notifiable disease requirements.
- ability to link all encounters for each individual both to the enrollment file and to provider characteristics included in administrative files, in order to distinguish incidence and prevalence.
- uniform, standardized definitions for data items and uniformed classification systems developed through a participatory consensus process.
- sufficient numbers of trained personnel in regional data centers to convert data into usable form for analysis, policy formulation, public health actions, and research.
- sufficient numbers of trained personnel in health alliances, state health departments, and Federal health agencies with skills in interpretation and analysis of network data in order to convert raw data into usable health information.
- automated monitoring of sentinel events and identification of deviations from expected numbers.

## **Enrollment Data**

For each individual who is eligible for benefits, information must be obtained to characterize that person through variables of public health importance. This information would most likely be obtained at the time of enrollment. It must be updated regularly, perhaps annually, to ensure that it is current, but, when appropriate, old entries should be preserved so that changes in variables of interest (e.g., body weight, occupation) can be studied.

CDC/ATSDR should advocate that at a minimum the following items should be included in an enrollment data set:

- patient's unique identification number
- identification of provider
- age, sex, and race/ethnicity
- occupation and industry, including occupational history
- marital status (e.g., single, married, widowed)
- numbers of children and other dependents
- unique identification numbers of other household members
- geographic location and zip code of residence, preferably down to the level of census tract
- educational attainment level
- date and place of birth
- risk-factor data (e.g., smoking status)
- data on socioeconomic status (e.g., household income)

## **Encounter Data**

For inpatient and outpatient encounters, CDC/ATSDR should advocate that a minimum data set established for each encounter include clinical detail sufficient to characterize the patient and his/her condition, including

- patient's unique identification number
- provider's unique identification number
- place of encounter
- reason for the visit
- diagnosis (multiple entries, if necessary)
- external cause of injury
- procedures (including laboratory tests) and results (including counseling/screening tests)
- therapies prescribed, administered, or discontinued
- disposition (e.g., return visit, follow-up, discharge, or death)
- total charges, as appropriate
- prevention interventions (e.g., counseling, screening tests, and health education)

## **Administrative Data**

CDC/ATSDR should advocate that administrative files made available by the plans should include

- number of participants enrolled.
- number of beds, examining rooms, operating rooms, or other areas for treatment or diagnostic procedures for all treatment facilities in which patients enrolled in the plans may be seen.
- numbers and types of preventive services offered and delivered by the Health Plan's providers at each facility.
- numbers and types of all health care personnel (e.g., board-certified neurologist, or medical/surgical registered nurses).
- geographic areas from which patients come.
- accounting of providers' relationships to plan (e.g., employee, independent contractor).
- information characterizing the areas in which Health Plans operate.
- information sufficient to uniquely identify all providers operating under the aegis of the plan.

## **Quality Management Program**

In order to provide incentives for preventive services, CDC/ATSDR should develop a list of data items to include in data systems developed for a Quality Management Program. Such a list should include information that public health officials deem necessary for decisions regarding prevention and intervention efforts.

## **Consumer Survey**

Consumer surveys would obtain information on consumers' satisfaction with health plans, access to care, and other items to be used in choosing among competing Health Plans (see Appendix A). CDC/ATSDR should advocate that a consumer survey include the followings:

- The scope and content must address more than patient satisfaction; analytic objectives beyond report card measures should be required.
- Sampling approaches should allow data to be aggregated by geographic areas such as community or state, not just by the area of the health plan.
- Sampling should be done in a way that allows adequate assessment of populations that are currently underserved or are particularly vulnerable to changes in the health system, as well as assessment of special populations of public health interest.
- The consumer survey must be integrated with other population health surveys to assure comparability, benchmarks, and appropriate adjustments for undercoverage.
- Information obtained through this survey should be linked to enrollment and encounter data in the network to facilitate meaningful analysis.

## APPENDIX E

### DEFINITIONS OF TERMS

#### Terms Related to Information

**data:** elemental descriptors or facts (e.g., sex=male, age=33 years)

**information:** data that are summarized, organized, or analyzed in a useful way

**knowledge:** information that has been synthesized, integrated, and interpreted into concepts

#### Terms Used in Computing and Telecommunications (Information Systems)

**back end:** software or hardware used to process data entered elsewhere or at a different time

*(Comment: May include computer programs to tabulate or otherwise analyze the data.)*

**back up:** periodic duplication and safe storage of data to prevent its loss

**communication:** the exchange of thoughts, data, information, or the like, via any medium (e.g., speech, written word, telecommunication)

**data base:** data stored in an organized and accessible manner

**data integrity:** protection of data from corruption, loss, destruction, or theft

*(Comment: Examples of procedures to ensure data integrity are routine backup of data stored on disks and use of passwords to limit access to data.)*

**front end:** software or hardware by means of which data enters a computer system

*(Comment: May include the preliminary processing of data at the time of data entry.)*

**gateway:** interface that operates under specific controls (e.g., a specific protocol)

**hardware:** physical components of a computer system

*(Comment: Examples of hardware are keyboard, central processing unit (CPU), monitor.)*

**interface:** software or hardware for transferring data from one computer system to another

*(Comment: Sometimes used to denote the software or hardware used to interface a computer system with the human user.)*

**link (verb):** to connect or couple data, data bases, data files, or any other data structures through the use of an element (e.g., name, Social Security Number, unique identifier) common to each data file or structure for the purpose of retrieval or analysis

**network:** a group or system of components that resembles a net in concept or form, with dispersed but interconnecting lines of communication  
(*Comment: This term often applies to methods of interconnecting computers.*)

**operating system:** set of instructions that coordinates the basic activities (copying and moving files) and parts (keyboard, monitor, disk storage) of a computer  
(*Comment: Operating system instructions may be contained in an integrated circuit e.g., ROM-BIOS or software [UNIX, DOS].*)

**smart card:** electronic file for storing individual health data or information for locating health data, with read and write capacity  
(*Comment: May have the capacity to process data.*)

**software:** a set of instructions to carry out specific activities on a computer  
(*Comment: Examples of software are WordPerfect, dBASE IV, Lotus 123.*)

## Terms Related to Surveillance

**assessment:** the determination of the importance, size, or value of  
(*Comment: Implies a quantitative aspect. Also an ordering or comparing. Assessment is performed at the level of knowledge or information, rather than at the level of data.*)

**monitoring:** ongoing, systematic collection and analysis of data for specific purpose  
(*Comment: Monitoring, unlike surveillance, does not include a dissemination component and may not be linked to program activities. In public health, its usual purpose is to detect changes in the environment or health status of populations.*)

**public health information system:** a system for storing, making available, and disseminating public health data, information, or knowledge  
(*Comment: Attributes include dissemination, ability to query, storage capability, locator or indexing. Can be the dissemination arm of surveillance. Can be queried by outsiders. Can contain information or knowledge not derived from surveillance.*)

**public health surveillance:** ongoing, systematic collection, analysis, and interpretation of outcome-specific, closely integrated with the timely dissemination of these data to those responsible for preventing and controlling disease.  
(*Comment: Surveillance data must above all be useful. They should be essential to the planning, implementation, and evaluation of programs and their activities.*)

**survey:** method of data collection for analysis of some aspect of a group or area  
*(Comment: Usually involves a sample of the target population, statistical or otherwise. Usually a one time collection, although it may repeat; snapshot in time; usually covers defined targeted participants. Does not necessarily include a specific dissemination component.)*

**system:** organized, interrelated collection or set of processes forming a unified whole

### Terms Related to Privacy

**confidentiality:** protection from unauthorized access, release, or dissemination  
*(Comment: Confidentiality is what is done to maintain/protect privacy. It is the obligation of the receiver or holder of information to maintain the privacy of the provider of the information.)*

**encryption:** technique or method of altering or transforming data so as to make them uninterpretable or unusable by anyone except persons with authorized access  
*(Comment: Done to maintain confidentiality in process of transmission, communication, or, in some cases, storage. In special cases, may even be one-way [e.g., creation of a unique identifier that cannot be decrypted to link with an individual].)*

**password:** a code that allows authorized access to data or a computer system  
*(Comment: Password protection is a term used for the method of using a password to protect the confidentiality of data. A physical barrier [such as locked storage of electronic storage media] is another technique or method that can be used to limited access to data.)*

**personal identifiers:** data that uniquely identify an individual  
*(Comment: Examples of personal identifiers are name and Social Security number.)*

**privacy:** freedom from unauthorized intrusion or access (e.g., to data or information)  
*(Comment: Privacy is the right of the provider of information.)*

**security:** protection of data and information resources to ensure continuity of services and the prevention of unauthorized access, loss, or unintended release of sensitive data or information.  
*(Comment: Often used to make data available only to those persons who should have access to them.)*

### Terms Used in Health Reform

**alliance:** consortium of health care providers and insurance plans organized and overseen by states, the purpose of which is to deliver health care services to defined populations in

defined geographic areas

**clinical care:** delivery of care to (and for) individuals, usually by individual practitioners

**data on consumers:** data gathered about a population concerning such characteristics as risk factors, health status, activities of daily living, and quality of life.

*(Comment: These data may be useful in anticipating the need for and cost of clinical care and public health services for the population.)*

**core data set:** Minimum set of data items to be obtained about each encounter a patient has with the health care system

**encounter data:** data associated with an interaction between a patient and the health care system (e.g., physician's office, clinical laboratory, radiology). The purpose of the interaction is screening, diagnosis, or treatment, principally for the benefit of the individual patient

*(Comment: May be used for public health surveillance.)*

**enrollment data:** data obtained at the time a person first enrolls in a health care plan. Data include information on demographic, administrative, risk-factor, financial, and past medical aspects of the patient and his/her care

*(Comment: Data may also be obtained or updated on a periodic basis following the initial visit.)*

**health care:** services provided to individuals or communities for the purpose of promoting, maintaining, monitoring, or restoring health

*(Comment: Health care includes clinical care and public health.)*

**provider:** any entity that delivers health care to any person

*(Comment: Provider can refer to any level of care, including individual physician, hospital, health maintenance organization, state-based system or alliance.)*

**data on provider:** data about the services and practices that health care providers make available in a given location or for a given population

*(Comment: These data may be useful for monitoring care and treatment practices.)*

**public health practice:** the applied science to ensure the health of a population rather than of individuals

*(Comment: Historically performed by government agencies rather than individual practitioners. May include prevention and treatment.)*

**report card:** a method or set of criteria used to evaluate differences in the performance of health care delivery organizations

*(Comment: Refers to a provider-specific evaluation of services/outcomes.)*

