Shaping a Health Statistics Vision for the 21st Century

Final Report
November 2002

Department of Health and Human Services Data Council
Centers for Disease Control and Prevention
National Center for Health Statistics
National Committee on Vital and Health Statistics

Table of Contents

Table of Contents . . . iii
Acknowledgments . . . iv
Foreword . . . v
Executive Summary . . . vi
Preface . . . xvi
Chapter One: What are Health Statistics? . . . 1
Defining health statistics . . . 1
Identifying the purposes of health statistics . . . 2
Uses of health statistics . . . 3
Defining the health statistics enterprise . . . 5
Chapter Two: A Model of the Influences on the Population's Health . . . 8
Need for and uses of the model . . . 8
Overview of the model . . . 8
Mapping the current enterprise using the model . . . 13
Using the model to describe issues and gaps . . . 15
Chapter Three: Developing, Evaluating and Improving Health Statistics . . . 20
The health statistics enterprise and the Health Statistics Cycle . . . 20
Elements in the Health Statistics Cycle . . . 20
Mapping the current system using the model of the Health Statistics Cycle . . . 26
Major issues and gaps in health statistics development, evaluation, and improvement . . . 28
Chapter Four: A Vision for the Health Statistics Enterprise . . . 35
Mission . . . 35
Overarching conceptual framework . . . 35
Core values . . . 36
Guiding principles . . . 36
The health statistics enterprise and the NHII . . . 47
NCVHS Recommendations for Achieving the Health Statistics Vision for the 21st Century . . . 49
Developing the recommendations . . . 49
Overview of recommendations . . . 49
Detailed recommendations for achieving the ten guiding principles . . . 50
Recommendations for moving toward the adoption of the overarching conceptual framework . . . 62
Summary of recommendations . . . 66
Glossary . . . 69
Table 1: Approaches for gathering information on influences on health . . . 72
Table 2: Programs and activities in the Health Statistics Cycle . . . 80
Appendix: National Committee on Vital and Health Statistics, 2002 Membership . . . 82

Acknowledgments

The development, writing, and production of this report involved a collaboration of many individuals and organizations, starting with the National Committee on Vital and Health Statistics (NCVHS), the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics (NCHS), and the DHHS Data Council. The principal authors of the report were Daniel J. Friedman (NCVHS), Edward L. Hunter (CDC/NCHS), and R. Gibson Parrish (CDC). The NCVHS Workgroup on Health Statistics for the 21st Century provided invaluable ongoing guidance, editing, and commentary: Vickie Mays, Paul Newacheck, and Barbara Starfield (Daniel J. Friedman, chair). The report also benefited immeasurably from the generous contributions of the many individuals who participated in the extensive national consultative process outlined in Figure 1 of the report.

Essential roles in the development of the report were also played by Robert Weinzimer, Marjorie Greenberg, Lisa Broitman, and Debbie Jackson who served as staff to NCVHS and NCHS; Susan Kanaan, who wrote the executive summary and the policy makers' summary; Judy Kaplan, who provided invaluable editorial advice; and Odell D. Eldridge, who coordinated the formatting and production of the report. Finally, Don Detmer and John Lumpkin played important roles in encouraging and supporting the process as Chairs of the NCVHS, as did NCHS Director, Edward J. Sondik, who initially conceptualized the need for such a re-evaluation of American health statistics.

Foreword

Information is the essential, invisible infrastructure upon which public health and health policy is built. Yet, the role of health information is poorly articulated, and the processes and mechanisms through which our nation manages information and translates it into effective policies and programs are poorly understood.

Two recent reports seek to redress this problem. First, in November 2001 the National Committee on Vital and Health Statistics (NCVHS) published *Information for Health: A Strategy for Building the National Health Information Infrastructure (NHII)*. The NHII report outlines strategic directions to share information and knowledge appropriately, when and where needed. The interconnections and utility of a functioning NHII can, with foresight, planning and vision, improve the public's health and health care.

Now, the attached report *Shaping a Health Statistics Vision for the 21st Century* describes a strategic vision for health statistics and the health statistics enterprise in the U.S. Health statistics are numerical data that characterize the health of the population and the influences and actions that affect the health of the population. Health statistics can be used to create fundamental knowledge and to guide health policy and programs. This report, developed through a partnership of the NCVHS, the Centers for Disease Control and Prevention's National Center for Health Statistics, and the Department of Health and Human Services (DHHS) Data Council draws on a national consultative process that involved numerous stakeholders in the health statistics enterprise. It also includes NCVHS recommendations for implementing this vision and calls for stronger integration and coordination in order to improve health statistics and ultimately the nation's health.

These two reports and their recommendations complement each other in important ways; some of these interconnections are described in the body of the attached report. The development of the NHII is essential for improving the health statistics enterprise and providing the electronic infrastructure through which data can be compiled, communicated, and translated into usable statistics. Both reports emphasize the importance of improved coordination of efforts, call for the development of standards to guide multiple parties, and stress the importance of strong protections for individually identifiable health information.

The tragic events in the country's recent history underscore the immediate need for a fully functioning and effective public health system, which will, in turn, rest on a strong health statistics enterprise and NHII. We hope that the principles and recommendations cited in this report, along with the recent report on the NHII, can prompt the strategic actions required to move forward toward this goal.

John R. Lumpkin, M.D., M.P.H.
Chair, National Committee on Vital and Health Statistics

Executive Summary

Preface

Health statistics are a cornerstone of our health system. They provide us with critical data to assess the health of our population and to make informed decisions about how to best direct our health-related resources and activities. However, health statistics in the United States have been hampered by a lack of clarity about definitions, appropriate foci, and agendas. Although health statistics already provide important information for health policy and programs, current efforts do not reflect a broad view of population health and the factors that influence health, nor coalesce around a shared vision that could bring about greater cohesion and efficiency in the development and use of health statistics. Because of this, health statistics have not achieved their full potential to support and guide policy and programs in the U.S.

To develop this shared vision for U.S. health statistics, three groups led a national consultation process that involved a broad range of stakeholders and lasted more than two years. The National Committee on Vital and Health Statistics (NCVHS)*, the National Center for Health Statistics (CDC/NCHS), and the Data Council of the U.S. Department of Health and Human Services (DHHS) offered stakeholders multiple opportunities to share their experience and contribute to the vision. The effort, which is summarized in Figure 1 on page xviii, was intended to provide broad-based guidance on the future of health statistics to federal, state and local governments, as well as to academia and the private sector.

More than 200 stakeholders—policy makers, public health and health care professionals, health statistics practitioners, researchers, and others—took advantage of the following opportunities to give input:

- Four expert discussion groups in Washington, D.C.
- Three local discussion groups in New Mexico, Pennsylvania, and Louisiana
- A National Academy of Sciences workshop
- Four public hearings in Chicago, Raleigh-Durham, San Francisco, and Washington, D.C.
- Presentations and town hall meetings at multiple national professional meetings
- Invitations to provide written comments

This document presents the results of the visioning process: a definition of health statistics, models of the factors that influence the health of populations and of an ideal health statistics cycle, and a vision for the health statistics enterprise. The document concludes with NCVHS recommendations for implementing the vision. The Committee's recommendations place special emphasis on building a stronger integrating and coordinating "hub" for the health statistics enterprise.

*NCVHS is the statutorily mandated public advisory body on health information policy to the Secretary of the U.S. Department of Health and Human Services. The NCVHS Workgroup on the 21st Century Vision for Health Statistics conducted this project for the Committee.

1. What Are Health Statistics

Health statistics are numerical data that characterize the health of a population and the factors that influence its health. In contrast with the related terms, "health data" and "health information," health statistics are distinguished by their focus on 1) quantification, 2) aggregation of data from observations on individuals, their communities, and the context of their communities, and 3) population health and the influences on it.

Health statistics are based on such information sources as surveys, patient encounters with the health care system, registries, and health administrative data. They facilitate understanding of the health status of individuals, groups, and society as a whole. Health statistics have three major uses: creating fundamental knowledge; guiding the assessment, development, and evaluation of health policy; and informing and evaluating the impact of population health programs and interventions as well as personal health decisions. They provide a basis for public and private decisions at local, state, and national levels.

The health statistics enterprise is the infrastructure and the activities or processes that produce health statistics. The highly decentralized and diverse health statistics infrastructure in the U.S. consists of public and private organizations and individuals at all geopolitical levels. Its activities include the collection, aggregation and compilation, analysis and translation, and evaluation of data. The health statistics enterprise in the U.S. has these characteristics:

- It focuses on the health of the entire U.S. population and its subpopulations.
- It serves the public interest and generates products that are public goods.
- It engages in systematic and organized inquiry.
- It strives for scientific objectivity.
- It involves multiple disciplines.
- It links U.S. health statistics activities to those of international partners in health and health statistics.

2. A Model of the Influences on the Population's Health

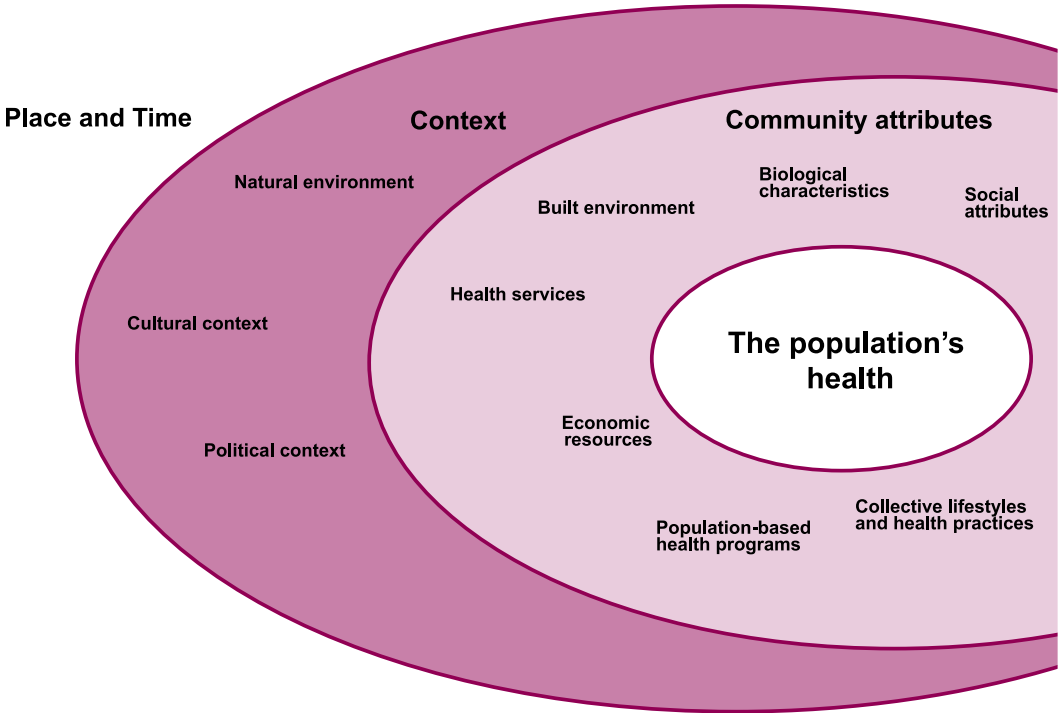
Many factors influence the health of a population, and to be useful, health statistics must provide a comprehensive and coherent picture of them all. Gathering and presenting data on diseases alone limits understanding of the complex interactions that affect health and encourages concentration on the prevention and management of disease instead of a more broadly integrated approach to maximizing health and reducing illness.

In the health statistics visioning process, the participants agreed on the need for a model with a broad view of population health and the influences on it. Such a model

could guide the regeneration, ongoing management, and evaluation of the health statistics enterprise. It could help identify data gaps and aid in understanding specific health issues.

Based on the consultative process and the work of several authors, a model (Figure 2 of the report, found in its full detail on page 9) was developed with three components:

- Measures of a population's health, i.e., of the entire U.S. population or any defined subpopulation
- Factors acting at community and individual levels (biological characteristics, the built environment, social and economic resources, health programs, collective lifestyles and health practices, and health services)
- The context or broad setting in which the population exists and acts (the natural environment, the cultural context, and the political context)



Influences on the population's health

This model of the influences on and measures of population health is a useful tool for organizing and describing the data gathered and analyzed by the health statistics enterprise at national, state, and local levels (as shown in Table 1 on page 72). It reveals the unevenness in the enterprise's capacity to provide data on all influences on population health. This unevenness is the result of a number of systemic problems, including:

- A failure to incorporate a broad definition of the population's health
- A predominant focus on individual health and its determinants
- A shortage of state and local data, especially survey data
- Inconsistent quality and quantity of data on vulnerable populations
- Inadequate longitudinal and life-cycle data
- A failure to adequately use available data and data sources

3. Developing, Evaluating and Improving Health Statistics

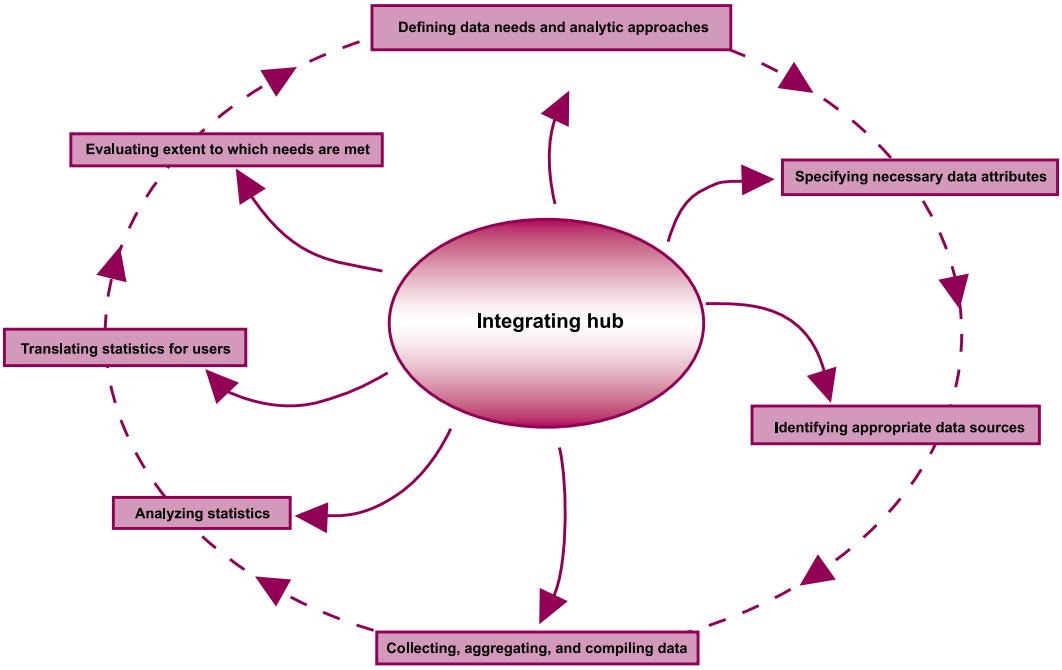
To support population health effectively, the health statistics enterprise must engage in a continuing process of evaluation, improvement, and regeneration. The stages of this process are part of a "health statistics cycle" (Figure 3 of the full report, found on page 21) that has seven components, ideally tied together by an integrating "hub." The components are:

- Defining data needs and analytic approaches
- Specifying necessary data attributes
- Identifying appropriate data sources and strategies
- Collecting, aggregating, and compiling data
- Analyzing statistics
- Translating statistics for users
- Evaluating the extent to which needs are met

Table 2 of the full report (page 80) shows the major elements of the health statistics cycle with examples of the organizations and entities that carry out its functions at national, state and local levels.

Although the health statistics enterprise has many strengths, and important data are collected, the vision process participants repeatedly pointed out major gaps in the availability of data and in the ability of the current health statistics enterprise to address data users' needs. The major causes of these gaps are:

- Insufficient connections between producers and users of data
- Lack of geographic and other detail
- Lack of timeliness in making data available
- Difficulty of finding and using existing data
- Lack of resources, jeopardizing major data sources
- Lack of enterprise-wide coordination and integration



Health statistics cycle

The hub that is needed to integrate the activities of the health statistics cycle should coordinate these activities, stimulate collaborative efforts, set standards, periodically review and revise the vision, generate new methods and approaches, and protect privacy, confidentiality and security. It was the consensus of participants that new approaches to enterprise-wide coordination and integration are needed because existing efforts are largely ineffective.

4. A Vision for the Health Statistics Enterprise

Out of the consultation and collaboration process emerged a vision for 21st century health statistics that was refined by NCVHS, CDC/NCHS, and the DHHS Data Council. The vision involves a mission, an overarching conceptual framework, core values, and guiding principles. This report proposes that the mission of the health statistics enterprise is *to efficiently provide timely, accurate, and relevant information that can be used to improve the population's health*. This can be achieved by operating within an overarching conceptual framework, adopting four core values, and implementing ten guiding principles.

The conceptual framework for the 21st century health statistics vision has these attributes:

- It focuses on health, the population, and the community.
- It emphasizes the distribution and level of health within the population.
- It delineates major influences on health.
- It guides a research agenda for improving the health of the population.

Briefly stated, the core values for the health statistics enterprise are:

- Maintaining the confidentiality and security of individually identifiable health information
- Maximizing the scientific integrity of all aspects of health statistics while acknowledging the specific ways in which the political, cultural, and business contexts may affect data collection, analysis, and interpretation
- Optimizing the enterprise's accountability to its users to ensure the availability of the information that is needed for improving the nation's health
- Ensuring the enterprise's accountability to its data suppliers to minimize their burden and to provide them with timely feedback.

Actualizing these core values requires the implementation of ten guiding principles:

1. *Enterprise-wide planning and coordination* to ensure relevance to local, state, and national policy and program decision-making and to an overall conceptual framework of the influences on the health of populations
2. *Broad collaboration* among data users, producers, and suppliers at local, state, and national levels to ensure efficiency of the health statistics enterprise and usefulness of the data that it produces
3. Rigorous policies and procedures for *protecting the privacy* of individuals and the confidentiality and security of data
4. *Flexibility* to identify and address emergent health issues and needs
5. *Use of data standards* to facilitate sharing and comparability of data
6. *Sufficient detail* at different levels of aggregation to support local, state, and national policy and programmatic decision-making
7. *Integrated, streamlined data collection* for multiple purposes
8. *Timely production* of valid and reliable health statistics
9. Appropriate *access to and ease of use* of health statistics
10. *Continuous evaluation* of the completeness, accuracy, and timeliness of health statistics and of the ability of the health statistics enterprise to support their production

NCVHS Recommendations for Achieving the Health Statistics Vision for the21st Century

After working closely with its partners to identify a vision for health statistics, the National Committee on Vital and Health Statistics developed recommendations for realizing the vision. The recommendations, which appear on pages 49-68 of the full report, are summarized in a table at the end of this Executive Summary.

NCVHS believes that the top priority for realizing the vision is to build the integrating and coordinating hub for the health statistics enterprise. Four NCVHS recommendations relate to this priority:

- Assign responsibility and authority for health statistics leadership within DHHS to a reconstituted National Center for Health Statistics, supported by a strong Board of Scientific Counselors.
- Establish a national Health Statistics Planning Board with overall responsibility for planning and agenda setting for health statistics in the United States.
- Designate overall responsibility for health statistics planning and coordination activities within each state's government to a single state agency, supported by autonomous state Health Statistics Planning Boards.
- Develop appropriate graduate and in-service training and continuous education addressing all elements of the health statistics cycle.

Carrying out these four recommendations, which relate to guiding principle 1, would facilitate the achievement of the Committee's recommendations for the other nine guiding principles. Those recommendations focus on research agendas, data set development, health statistics reports, dissemination of best practices, training, and other key topics. To ensure that health policy-making keeps its focus on the core influences on the health of populations, an overarching conceptual framework must guide the enterprise.

Finally, the 21st century health statistics enterprise in the U.S. must be viewed within the context of the National Health Information Infrastructure (NHII), which has the potential to connect all health decision-makers to sound information and to each other. NCVHS recommended a strategy for building the NHII in a November 2001 report.* The vision for health statistics contributes to the development of the Population Health Dimension of the NHII. Participants in the health statistics enterprise must become actively and enthusiastically engaged in this wider effort.

* NCVHS, *Information for Health: A Strategy for Building the National Health Information Infrastructure,* 15 November 2001.

Summary of NCVHS recommendations

Recommendations on data access

- Use the World Wide Web to improve access to national, state, and local data on the population's health and influences on the population's health. [GP 9.1]
- Disseminate throughout the health statistics enterprise the best methods for maximizing appropriate access to and ease of use of health statistics. [GP 9.3]
- Enhance regularly released health statistics reports to reflect an overarching conceptual framework of influences on the population's health. [CF 4]

Recommendations on data set development

- Develop systems to actively monitor the population's health and potential influences on the population's health in order to identify emerging problems. [GP 4.1]
- Develop mechanisms to rapidly modify the samples, data collection strategies, and data collection instruments of ongoing data collection systems in response to emerging issues. [GP 4.2]
- Geocode all ongoing data sets that feed the health statistics enterprise to the census block group level. [GP 6.2]
- Develop and validate topic-specific question modules for population-based surveys to support state and local agencies' development of small-area health statistics. [GP 6.3]
- Identify duplicative data collection efforts and eliminate redundancy. [GP 7.1]
- Support and fund efforts within individual states to integrate or better coordinate, as appropriate, discrete data systems. [GP 7.2]
- Support and fund ongoing multi-purpose data collection systems and data integration efforts. [GP 7.3]
- Develop and adopt benchmarks for timeliness of release of data and reports derived from those data. [GP 8.1]
- Assure that appropriate measures of functional status and well-being are included in ongoing systems that are a part of the health statistics enterprise. [CF 1]
- Prepare and provide to staff of state and Federal agencies with health statistics responsibilities a manual and a "tool kit" of data sets for use in community-level analyses. [CF 2]
- Develop person-based, longitudinal data sets and surveys in order to develop portraits of influences on the population's health throughout the life cycle. [CF 3]

Recommendations on data standards

- Harmonize the CDC public health conceptual data model with the overarching conceptual framework for the health statistics enterprise in order to facilitate the development of standards. [GP 5.1]
- Adopt or, if necessary, develop standards for data elements commonly used in all methods of data collection, for electronically transmitting data, for presenting and disseminating data, and for providing electronic access to data. [GP 5.2]
- Provide guidance for implementing and auditing the use of standards by the health statistics enterprise. [GP 5.3]
- Contribute to the national standards setting process. [GP 5.4]

Recommendations on enterprise structure

- Assign overall responsibility for health statistics leadership within DHHS to a reconstituted National Center for Health Statistics, supported by a strong Board of Scientific Counselors. [GP 1.1]
- Establish a national *Health Statistics Planning Board* with overall responsibility for planning and agenda setting for health statistics in the United States. [GP 1.2]
- Assign overall responsibility for health statistics planning and coordination activities within each state's government to a single state agency. Establish a *Health Statistics Planning Board* in each state with overall responsibility for planning and agenda setting for health statistics within the state. [GP 1.3]

Recommendations on evaluation

- Compile feedback on the practical applications of health statistics to policy, research, and public health practice and disseminate these findings to providers of data, policymakers, and the public. [GP 10.1]
- Conduct periodic evaluations of NCHS' adherence to the National Academy of Sciences' attributes of a Federal statistical agency. [GP 10.2]

Recommendations on privacy policy

- Develop a code of fair information practices for the health statistics enterprise. [GP 3.1]
- Develop guidance on the application of the HIPAA Standards for Privacy of Individually Identifiable Health Information to health statistics. [GP 3.2]
- Develop model state and Federal privacy, confidentiality, and security statutes for government-sponsored health statistics. [GP 3.3]
- Develop a toolbox of privacy, confidentiality, and security best practices for use throughout the health statistics enterprise. [GP 3.4]

Recommendations on a research agenda

- Develop and fund a research agenda for the study of methods and technologies to enhance the privacy, confidentiality, and security of health statistics. [GP 3.6]
- Develop and fund a research agenda to explore new data collection strategies that can rapidly and flexibly provide data on emerging influences on the population's health; assess the validity and reliability of items used in key ongoing data collection systems; and estimate any loss in accuracy from early publication of provisional, incomplete data from ongoing data collection systems. [GP 4.3]
- Develop methods to validly and reliably estimate important indicators of the health, and of the influences on the health, of state and local populations. [GP 6.1]
- Develop and implement methods and technologies that will maximize the timeliness of release from ongoing data collection systems of regularly scheduled reports. [GP 8.2]
- Provide controlled access to potentially identifiable individual-level data on the population's health and influences on the population's health. Such access would be provided, at designated sites, to researchers pursuing approved projects. [GP 9.2]
- Support research into how health statistics can be used more effectively to understand the causal relationships between specific influences on health and functionality and well-being in general and specific high priority diseases. [CF 6]

Recommendations on training

- Expanded graduate, in-service, and continuous training opportunities should be developed and offered in focusing on all elements of the health statistics enterprise. [GP 1.4]
- Develop a model training agenda and training materials for use throughout the health statistics enterprise focusing on fair information practices, privacy, confidentiality, and security. [GP 3.5]
- Provide training about concepts related to the population's health and influences on the population's health at professional meetings attended by staff of state and Federal agencies with health statistics responsibilities. [CF 5]

Preface

In 1998, Edward J. Sondik, Ph.D., Director of CDC's National Center for Health Statistics (NCHS), asked the National Committee on Vital and Health Statistics (NCVHS), the statutorily mandated public advisory body on health information policy to the Secretary of the U.S. Department of Health and Human Services (DHHS), to provide guidance on the development of goals for health statistics for the next ten to fifteen years. NCVHS responded by establishing the Workgroup on the 21st Century Vision for Health Statistics[1] and by forming a partnership among the Workgroup, NCHS, and the DHHS Data Council.[2]

In 1999, the NCVHS/NCHS/DHHS Data Council partnership initiated a 21st century health statistics vision process, which has included both fact-gathering and consensus-building. Through this process, the three partners attempted to identify groups and individuals that should become stakeholders in the vision process; to define health statistics; to delineate health statistics needs; and to develop a shared vision for health statistics in the U.S. This process has involved more than 200 participants, who have provided reflections, suggestions, and recommendations through national and local discussion groups, a National Academy of Sciences workshop, regional public hearings, and presentations and town hall sessions at meetings of professional associations.[3] Figure 1 provides an overview of the health statistics vision process from 1999 through 2002.

Opportunities for providing input into the vision process

- Four expert discussion groups held in Washington, D.C.–48 participants in total
- Three local discussion groups–held in Albuquerque, Harrisburg, and New Orleans–40 participants in total
- Workshop held in Washington, D.C., by the National Academy of Sciences' Committee on National Statistics–90 participants
- Four regional public hearings–held in metropolitan Chicago, Raleigh-Durham, San Francisco, and Washington, D.C.–74 participants in total
- Presentations and town hall meetings held at meetings of the Association for Health Services Research, the Association of Maternal and Child Health Programs, and the National Association for Public Health Statistics and Information Systems

[1]The Appendix contains a list of NCVHS members.

[2]The DHHS Data Council coordinates all data collection and analysis activities of DHHS. Its membership consists of all Assistant Secretary and Agency Administrator level DHHS officials, the DHHS Privacy Advocate, and the Senior Advisor to the Secretary on Health Statistics, a dual position held by the Director of NCHS.

[3]A record of these sessions and hearings, along with lists of many of the individuals making contributions to the vision process can be found at http://ncvhs.hhs.gov/hsvision/visiondevelopment.html [accessed 1 November 2002].

Participants in the vision process have generously committed their time and expertise. They offered detailed written and oral feedback on the June 2000 interim report[4] and were invited to proffer recommendations for inclusion in the report.

The health statistics vision process has revealed that health statistics in the United States have been hampered by lack of clarity about what constitute health statistics, about what practitioners of health statistics should focus on now, and about what they should focus on in the future. Chapter One defines health statistics and provides examples to illustrate their use and importance. Chapters Two and Three discuss the factors that influence the health of populations, the components of the health statistics enterprise in the United States that attempt to measure these factors, and the gaps in the enterprise. Following these background chapters, Chapter Four portrays a vision for the U.S. health statistics enterprise for the 21st century, including its mission, core values, and guiding principles. Chapters One through Four are joint products of the NCVHS/NCHS/DHHS Data Council partnership. Finally, the last section of the document presents NCVHS's recommendations for implementing the vision delineated in Chapter Four.

[4]NCVHS, Shaping a Vision for 21st Century Health Statistics: Interim Report, June 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/Vision21stCentury.pdf [accessed 1 September 2002].

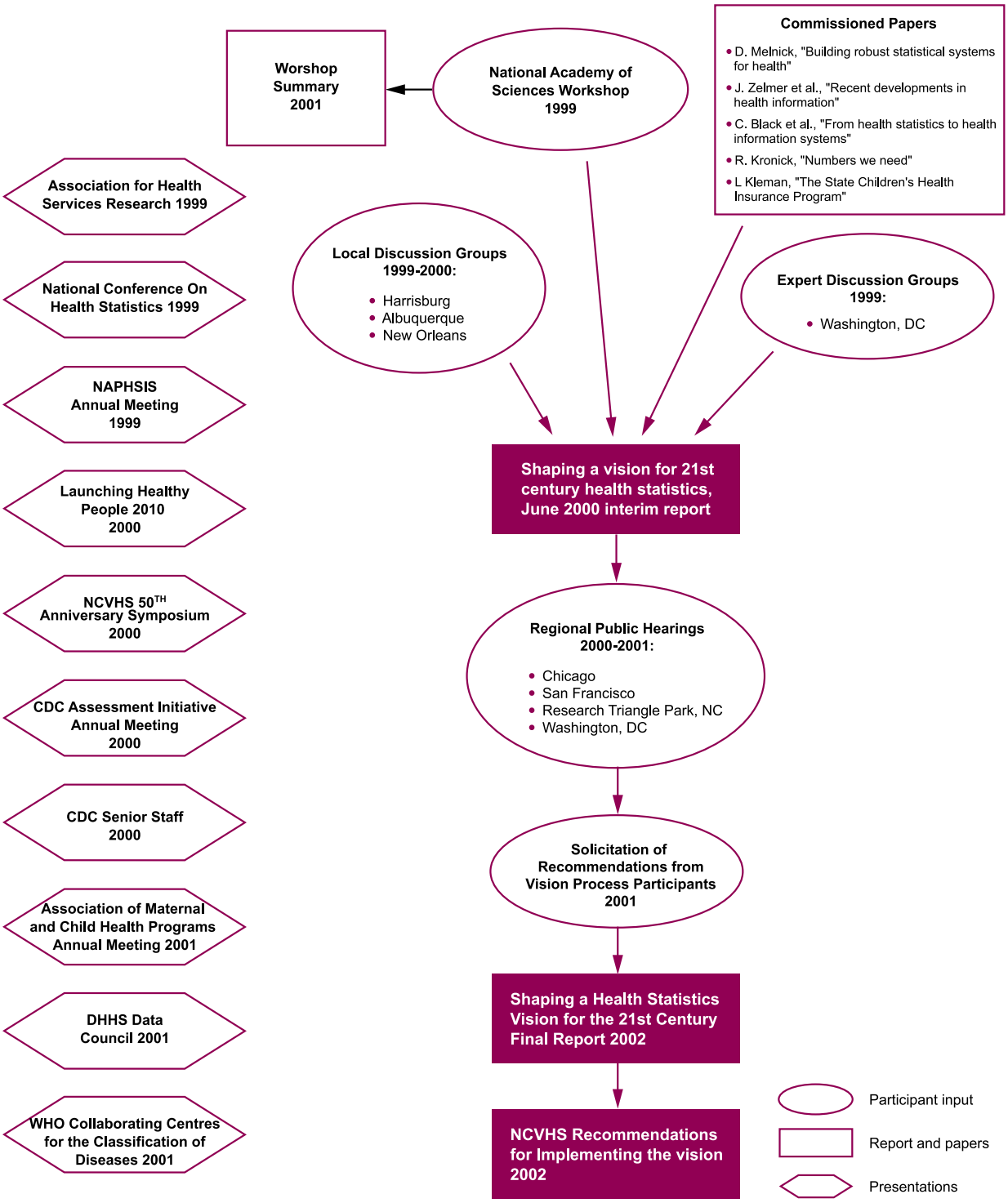


Figure 1. Joint NCVHS, NCHS, and DHHS Data Council process for developing the 21st century vision for health statistics

Chapter One: What are Health Statistics?

Defining health statistics

The collection, interpretation, and use of health statistics in the United States have been hampered by lack of clarity about what is included in "health statistics" and what the nation's health statistics priorities should be now and in the future. This lack of clarity is illustrated by the absence of definitions of health statistics in the American public health literature.[1]

Health statistics are defined in this report as numerical data that characterize the health of a population and the influences and actions that affect the health of a population. These influences include the environment; genetic and other biological characteristics; health services; community attributes; and the political and cultural contexts. Health statistics support the study of the interaction of these influences as well as the study of the way specific elements in each of these areas (such as the health services system or environmental exposures) influence the health of populations. Health statistics are used to design, implement, monitor, and evaluate specific health policies and programs. Properly organized and communicated, health statistics enable citizens, policy makers, public health workers, and health care providers to assess local or national health, mobilize to improve it, and evaluate the success of those efforts.

> Health statistics are numerical data that characterize the health of a population and the influences and actions that affect the health of a population.

Defining health statistics is difficult because there are other terms that are often used interchangeably but have other meanings or connotations. For example, the term "health data" is often used to refer both to a single factual observation (such as the age of an individual) and to the aggregation of such observations (such as the age distribution of a population). Health statistics, which focuses on populations, would only be used to refer to the latter.

Similarly, "health information" is very inclusive and may refer to raw, analyzed, quantitative, or qualitative observations, records, or other facts on individuals, groups of individuals, or populations. For example, health information refers not only to the age of an individual or to the age distribution of a population, but also to knowledge derived from research on the health effects of aging, and to patient-oriented information on how to treat illnesses associated with aging. Although health information includes health

[1] In fact, the only definition of health statistics in a frequently cited public health volume is in John Last's A Dictionary of Epidemiology (Last, J., ed., A Dictionary of Epidemiology, 4th edition, Oxford University Press, Oxford and New York, 2001). It is worthy of note that Professor Last is on the faculty of the University of Ottawa.

statistics, its broad and multiple uses render it overly general for the purpose of this report.[2]

Health statistics, then, are distinguished by their focus on 1) quantification, 2) aggregation of data from observations on individuals, their communities, and the context of their communities, and 3) population health and the influences on it.

Health statistics provide information about aggregations of people, institutions, organizations, or health events, rather than information about an individual person or event. Health statistics are created at the aggregate population or sub-population levels. However, health statistics are typically created from raw data collected on individuals (or from the context in which they live and work), and data linkage may occur at either geographic area or individual levels. Regardless of whether linkage occurs at the individual or aggregate level, health statistics provide information about populations and sub-populations, and never about individuals.

Gold standard data sets that are the source of health statistics contain information that can be scientifically generalized to the entire population or a designated sub-population. These gold standard data sets can also be evaluated for completeness of ascertainment. In practice, since gold standard data sets are not always available or affordable, specific data needs must be evaluated in order to determine the level of completeness and quality required.

Health statistics are based upon data sets derived from numerous data sources and collected through several different modes of data collection. Health statistics are drawn from population-based health surveys (such as the Behavioral Risk Factor Surveillance System); surveys of health service institutions and individual health care providers (such as the National Health Care Survey); administrative data generated through ongoing health service operations (such as Medicaid claims data or state-wide hospital discharge data); reportable disease registries (such as state participants in the National Program of Cancer Registries); vital records registries; and other data collection systems not directly focusing on health. Health statistics may be based upon data originally intended to characterize the health of a population and the factors that influence it (such as that collected in the National Health Interview Survey) or may be based upon data originally intended for other purposes (such as information collected through the Census of Population or the Environmental Protection Agency's Aerometric Information Retrieval System). The data sets upon which health statistics are based may be collected on a periodic or an ongoing basis.

Identifying the purposes of health statistics

Health statistics fulfill essential functions for public health, the health services system, and our society. Most basically, health statistics enable understanding of where

[2]Another term in common use is "surveillance." As applied to health, "surveillance is a continuous and systematic process of collection, analysis, interpretation, and dissemination of descriptive information for monitoring health problems (Buehler, J., "Surveillance," *Modern Epidemiology*, Rothman, K. and Greenland, S., ed., 2nd ed., Lipincott Williams & Wilkins, Philadelphia, 1998, pp. 435-457). The distinction between statistics and surveillance is the continuous nature of surveillance, its use of any descriptive information–not just numerical data, and its focus on health problems, as opposed to populations.

we stand in terms of health as individuals, as subgroups, and as a society. Through health statistics, we gain a collective understanding of our health, our collective experience with the health services system, and our public health problems and challenges. Health statistics establish a basis for comparisons between subpopulation groups or geographic areas. Health statistics enable us to look at the distribution of health in the U.S., revealing, for example, the existence of health disparities between racial, ethnic, and socioeconomic groups, between Americans with and without substantial functional disabilities, and across rural, suburban, and central city areas. Health statistics document our current and past reality, and provide us with the ability to identify health trends and to anticipate future trends in health and health services.

Health statistics provide us with the information upon which we can base important public decisions at the local, state, and national levels. Once we have made those public decisions, health statistics make us accountable for the decisions that we have made. Health statistics thus enable us to evaluate the impact of health policies and health programs on the public's health.

In short, health statistics give us the information we need to improve the population's health and to reduce health disparities.

Uses of health statistics

Health statistics have three major uses. *First use: creating fundamental knowledge about the health of populations and sub-populations, influences on health, and interactions among those influences.* The fundamental knowledge created by health statistics is varied, and can contribute to increasing our understanding of the health care system and the health of populations; of the influences on well-being, functional status, and disease; of relationships among community resources and health; and of changes in the health of a population and its health care, particularly as major changes are occurring in private markets and Federal and state policy.

Uses of health statistics

- Create fundamental knowledge
- Develop information to guide health policy development, assessment, and evaluation
- Generate information to guide implementation, targeting, evaluation, and refinement of population health programs and interventions and to guide personal health decisions

Use of health statistics to create fundamental knowledge: heart disease

Health statistics provide insight into factors associated with heart disease and its impact on the population. Large-scale population surveys yielded some of the first evidence of the association between heart disease and cholesterol, hypertension, and behaviors such as diet, exercise, and smoking. Statistics document the distribution of risk factors among the population, allowing us to understand differentials in health. And health statistics provide insight into the impact of heart disease on the population and economy through mortality data that track trends, surveys that document lost productive time to heart disease, and studies that document the burden of illness on the health care system and the economy.

Second use: developing information to guide health policy development, assessment, and evaluation. Health statistics help to establish health policy priorities, based upon assessment of the health of a population and its health trends, disparities, and needs as well as the quality and efficiency of health services. Health statistics are used to project the impact of alternative policy choices on a population's health and to measure change associated with policy implementation.

Use of health statistics to inform policy: health insurance

Health statistics provide insights into the health insurance coverage of the population, the differences in coverage between subpopulation groups (such as between the poor and the non-poor, the employed and the unemployed), and the relationship between insurance, access to care, and population health. Health statistics help set the health policy agenda and are closely watched for their implications for publicly funded health insurance programs. The Federal government alone spends more than $400 billion annually on health, primarily for its Medicare, Medicaid, and S-CHIP programs. Improving health statistics can have real returns in the efficiency and effectiveness of these and other programs.

Third use: generating information to guide implementation, targeting, evaluation, and refinement of health programs and other interventions for populations and to guide personal health decisions. Once policy directions have been established, health statistics are used to design public health programs that implement those policies. Health statistics enable identification and understanding of populations at risk of poor health and targeting of public health programs to those groups most in need of them. Once programs have been implemented, health statistics are used to evaluate them by determining their impact

on target groups. Health statistics also enable information to be made available to inform personal decisions about health, helping individuals understand health promotion and disease prevention strategies, the consequences of health risk behaviors, and health care outcomes associated with personal health care decisions.

Use of health statistics in health programs and interventions: asthma[3]

Health statistics have played a critical role in improving our understanding of, and our attempts to prevent and treat, asthma. Health statistics have shown that the prevalence of asthma has increased in the United States, as well as in many other developed countries over the past 20 years. They have also provided insight into what groups are most affected by asthma and, thereby, facilitated the targeting of programs to reduce the prevalence and adverse effects of asthma. For example, health statistics have shown that mortality is higher among African Americans and those of Puerto Rican heritage than among white non-Hispanics, and specific programs have been developed to help these groups.

These three uses of health statistics must guide and inform both the structure and the processes of the health statistics enterprise.

Defining the health statistics enterprise

The health statistics enterprise consists of the infrastructure and the activities or processes necessary to produce health statistics. The infrastructure of the health statistics enterprise consists of the public and private organizations and individuals at all geopolitical levels that perform the processes (that is, carry out the activities) of health statistics. The infrastructure of the U.S. health statistics enterprise is highly decentralized and diverse, without the overall organization and direction that often characterize corporate enterprises. The infrastructure of the health statistics enterprise includes the many organizations that collect, analyze, and disseminate data on the health of populations and on the factors that influence health: 1) Federal agencies, such as CDC/NCHS, the Center for Cost and Financing Studies of the Agency for Healthcare Research and Quality, and the Office of Applied Studies of the Substance Abuse and Mental Health Services Administration; 2) national organizations, such as the American Hospital Association; 3) state and local public agencies, such as state centers for health statistics within state health departments, and county public health departments; 4) private third-party payers, such as health maintenance organizations and other health plans; and 5) foundations and their grantees.

[3]References for asthma example:

a) "*Surveillance for asthma--United States, 1960-1995.*" Morb Mortal Wkly Rep. 1998;47(SS1):1-28.

b) Homa DM, Mannino DM, Lara M. "*Asthma mortality in US Hispanics of Mexican, Puerto Rican, and Cuban heritage: 1990-1995*." Am J Respir Crit Care Med. 2000;161:504-509.

c) *"Self-Reported asthma prevalence among adults ---United States, 2000."* Morb Mortal Wkly Rep. 2001;50(32):682.

d) *"Asthma hospitalizations and readmissions among children and young adults -- Wisconsin, 1991-1995."* Morb Mortal Wkly Rep.1997;46(31):726.

The infrastructure of the health statistics enterprise carries out many activities or processes that yield health statistics. As explained in greater detail in Chapter Three, the activities of the health statistics enterprise include the collection of data from institutions, organizations, and individuals; aggregation and compilation of those data into health statistics; analysis and translation of health statistics in order to make them comprehensible to varied users; and evaluation of health statistics and the health statistics enterprise in order to improve them.

The health statistics enterprise, as currently constituted in the U.S., can be described by the following major characteristics.

The U.S. health statistics enterprise

1. Focuses on the health of the population and subpopulations
2. Serves the public interest and generates products that are public goods
3. Engages in systematic and organized inquiry
4. Strives for scientific objectivity
5. Involves multiple disciplines
6. Links U.S. health statistics activities to those of international partners in health and health statistics

1. Focuses on the health of the population and the influences and actions that affect the health of the population. While different organizations, individuals, and processes within the health statistics enterprise may focus on different influences and actions, ultimately all parts of the enterprise contribute to better understanding and characterizing the population's and sub-populations' health.

2. Serves the public interest and generates products that are public goods. The health statistics enterprise includes both public and private sector entities, all of which can contribute to collection, aggregation, analysis, and translation of data. Government agencies, private sector health care providers, insurers, and purchasers develop health statistics. Each of the organizations and individuals comprising the health statistics enterprise ultimately contribute to better understanding and characterizing the population's health. Data made available in the public domain by these entities become public goods in the economic sense, in that once made available multiple users can use them for multiple purposes.

3. Engages in systematic and organized inquiry. The health statistics enterprise relies upon systematic data collection, aggregation, analysis, and translation. Health statistics entail organized inquiry and not unsystematic collection of anecdotes and personal impressions.

4. Strives for scientific objectivity. The health statistics enterprise strives for scientific objectivity in data collection, aggregation, analysis, and translation, and attempts to minimize or eliminate any explicit political biases in its activities. The choice of priorities within the health statistics enterprise is necessarily based upon a series of choices about what constitutes and how to measure the health of a population, and what influences on health merit study and how to assess their influences.[4] Such choices are inherently based upon values and judgments; health statistics practitioners make scientific objectivity a preeminent concern.

5. Involves multiple disciplines. The practice of health statistics necessarily involves multiple disciplines, including statistics, epidemiology, health services research, demography, public health informatics, information technology, and economics.

6. Links U.S. efforts to those of international partners in health and health statistics. The U.S. is guided by international health statistics standards and policies, such as periodic revisions to the International Classification of Diseases codes. The U.S. health statistics enterprise also contributes to international discussions of health statistics methods and provides health statistics to international agencies for comparative purposes.

[4]Krieger, N., "The Making of Public Health Data: Paradigms, Politics, and Policy," *Journal of Public Health Policy,* 1992, pp. 412-427.

Chapter Two: A Model of the Influences on the Population's Health

Need for and uses of the model

Many factors influence the health of a population. In order to develop useful health statistics that accurately characterize the health of a population and the influences and actions that affect the health of a population, it is essential to have a comprehensive and coherent representation of these factors. Through the extensive health statistics vision consultative process, clear agreement emerged on the need for a model that takes a broad view of the health of populations and the factors that influence their health. This model should guide the development, ongoing management, and evaluation of the health statistics enterprise so that it can provide data that are comprehensive and useful for characterizing health at all geopolitical levels. When applied to a population of interest at any geopolitical level for a specific time period, this model can serve to identify gaps in data for the given population.[5] In addition to being used to identify gaps, the model can be applied to a specific health issue (such as lung cancer, homicide, or depression) in order to identify the principal factors influencing it and potential approaches to preventing or ameliorating it.

Figure 2 presents a model of the factors that influence health that grew out of the health statistics vision consultative process and the work of several authors.[6] This chapter describes this model and uses it to evaluate the gaps in the health statistics currently available in the U.S. In Chapter Four, this model is used to guide the presentation of the vision for health statistics in the 21st century. Definitions of key terms in Figure 2 are included in the Glossary.

Overview of the model

The model has three components: 1) the context or broad setting in which the population exists and acts, 2) factors acting at the community and individual levels, and 3) measures of a population's health. The model positions a population's health as the central outcome variable (central oval in Figure 2). The health of a population is described by measures of disease, functional status, and well-being that reflect both the level and distribution of health in the population.

[5]Black, C., Roos, N., Roos, L., "From health statistics to health information systems: a new path for the 21st century, Available from http://www.ncvhs.hhs.gov/hsvision/CP--Black.pdf [Accessed 18 October 2002].

[6]For example, see: Lalonde, M., *A New Perspective on the Health of Canadians: A Working Document*, Government of Canada, Ottawa, 1974; Evans, R. and Stoddart, G., "Producing health, consuming health care," *Why Are Some People Healthy and Others Not?* Evans, R., Morris L. Barer, and Marmor, T., ed., Aldine de Gruyter, New York, 1994; Starfield, B., *Primary Care: Balancing Health Needs, Services, and Technology*, Oxford University Press, Oxford and New York, 1998; Berkman, L. and Glass, T., "Social integration, social networks, social support, and health," *Social Epidemiology*, Berkman, L. and Kawachi, I., ed., Oxford University Press, Oxford and New York, 2000; and U.S. Department of Health and Human Services, *Healthy People 2010: Understanding and Improving Health*, U.S. Department of Health and Human Services, Washington, D.C., 2000.

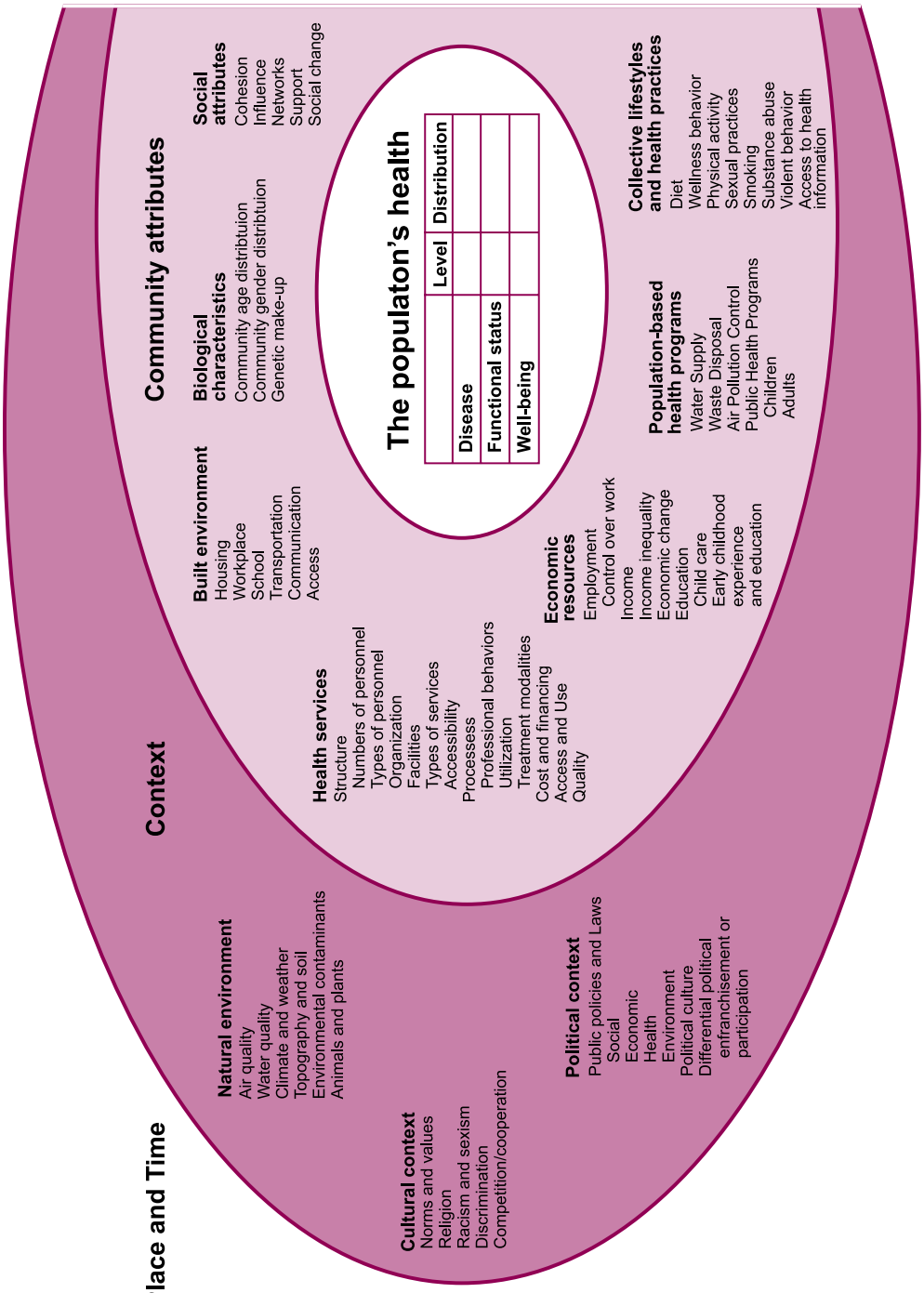


Figure 2. Influences on the Population's Health

All variables in this model are either "aggregate" or "ecological." Aggregate measures represent those community attributes that are derivable from the attributes of individual members of the community, such as the community age structure or the incidence rate for sexually transmitted diseases. Aggregate measures represent either averaged data on behaviors or attributes (such as the mean or median household income or the poverty rate) or a function of the distribution of behaviors or attributes (such as the ratio of the mean or median household income of the lowest fifth in the income distribution to the mean or median household income of the highest fifth). Ecological measures represent those community attributes that are not derivable from the attributes of individual members of the community. Levels of air or water pollutants, public policies, or the structure of health services are examples of ecological variables.

Social attributes, biological characteristics, the built environment, health services, economic factors, population-based health programs, and collective lifestyles and health practices (the major categories in the "community attributes" band immediately surrounding the population health oval in Figure 2) are factors that immediately affect the health of a population. Depending upon the specific health issue, these factors may interact in different ways, and may influence the population's health in different ways. The built environment, health services, social attributes, and health programs for populations function predominantly as ecological variables in their influences on health. Biological characteristics, individual or family economic resources, and collective lifestyles and health practices function predominantly as aggregate variables.

The natural, cultural, and political contexts (the major categories within the outermost, "context", band in Figure 2) are ecological variables that affect both the population's health and the health of individuals.

Finally, it is important to recognize that the influences portrayed in Figure 2 must be used to describe a particular place and time.[7, 8] The configuration of the influences on the population's health will differ from place to place (in other words, from geopolitical area to geopolitical area) and will also differ depending upon the particular time frame (in other words, the year or historic era). The location of a particular population in place and time will determine the specific constellation of contextual influences and community resource influences that most strongly affect the population's health.

The model does not presume causality, directionality, or interactions among its components. As indicated above, the purpose of Figure 2 is to provide a picture of the influences on a population's health for use in evaluating the availability of data and gaps in data at the national, state, and or local levels. A variety of research hypotheses can be developed to test the existence and direction of causality among elements in the model. Similarly, the model does not posit the relative weight or importance of different

[7]Hayes, M., "Evidence, determinants of health and population epidemiology: humming the tune, learning the lyrics," *The Determinants of Population Health: A Critical Assessment*, Hayes, M., Foster, L., and Foster, H., ed., Department of Geography, University of Victoria, Victoria, British Columbia, 1994.

[8]Hayes, M. and Dunn, J., *Population health in Canada: a systematic review*, Canadian Policy Research Networks, Inc., Ottawa, 1998.

influences or combinations of influences. It leaves this work to analysts and researchers. The contribution of the health statistics enterprise is to ensure the availability of data to support such analysis and research.

The next section of this chapter provides a more detailed description of the elements within each of the three components of the model.

Model Overview – Context

The "context" band consists of three major elements: the natural environment, the cultural context, and the political context. The natural environment has a strong overall influence on the population. Climate, topography, and water resources affect the animals and plants that can live in the area. These, in turn, affect food resources and the size and density of the populations that can be supported. Proximity to waterways, seas, and oceans and topography determine trade routes and the extent of interaction with other peoples.[9] All of these influence health. The cultural context refers to the norms and values of a population; these, in turn affect the level of competition and cooperation both within a population and between the population and other neighboring populations. The presence and extent of advantages held by or discrimination against particular population groups are manifestations of norms and values that can affect the health services utilization, health practices, and health of these groups.

The political context is the expression of norms and values in the overall political culture of a given population and geopolitical area. A society codifies its norms and values as the public policies and laws that guide the actions of its members. Thus, the political context can significantly shape the society's impact on the natural environment and how the society functions socially and economically. Policies and laws may also determine the availability and nature of health programs for the population and of health services. Finally, the extent to which individuals and groups within the population are enfranchised and participate in the political culture of a society has a strong influence on the responsiveness of the society to their health needs and other needs.

Model Overview – Community Attributes

Community attributes exist within and are determined by their environmental, cultural, and political context and the population that shares in this context. These attributes include the community's social attributes, biological characteristics of the population, the built environment, health services, economic resources, population-based health programs, and collective lifestyles and health practices of people in the community.

The social attributes of the community consist of the type and extent of social networks, the overall social cohesion of the community, and the support provided by the

[9]Diamond, J., *Guns, Germs, and Steel: The Fates of Human Societies.* New York, W. W. Norton, 1999.

community or by subpopulations within the community to its members. Discrimination against particular groups within the community may adversely affect the health of these groups by excluding them from important social networks and support.

The community's biological characteristics include its age and gender structures, the genetic make-up of its membership (such as the prevalence in the population of genes associated with hemoglobinopathy or hemochromatosis), and the immune status of its membership (such as the level of immunity to measles as a result of widespread childhood vaccination or infection). The community age structure affects the health outcomes of the population; for example, communities with relatively older populations are likely to have higher prevalences of chronic diseases.

The built environment refers to modifications of the natural environment to support human habitation and activity, such as housing, roads and canals, telecommunications infrastructure, work places, and schools. These have a profound influence on the health of the community by determining the level of protection from weather and climate and the ability of the community to interact and trade with other communities. The quality of the physical structures available for schooling and work also influence the nature and quality of these activities. Some of these structures also provide the setting for important familial, community, and other social interactions (for example, housing, schools, and work places).

The economic attributes of the community consist of the overall wealth of the community and its members, the distribution of this wealth, and the type, extent, and stability of employment of its members. As with social attributes, the extent of equity in the distribution of, and access to, economic attributes for individuals and groups is an important influence on community health. Furthermore, the control that members of the community can exercise over their work and economic status can profoundly affect health. Finally, educational opportunities for children and adults, and the level of education of members of the community are both important influences on, and markers of, the economic and health status of the community.[10]

The cultural and political contexts and the social and economic attributes of a community exert a profound effect on the type, availability, and quality of health services and health programs for populations. The overall effectiveness or quality of health services depends on their structure and capacity (including factors such as the number, type, and training of personnel; the quality and availability of facilities; the method and level of financing), their processes (organization of the delivery system, professional behaviors, and utilization), and accessibility.

In addition to health services, which are usually delivered on an individual basis, population-based health programs and the provision of health-related information to the community exert important influences on health. Population-based health programs

[10]It has been suggested that child development be included as an influence on health. (Hertzman, C., "The case for child development as a determinant of health," *Can J Public Health*, 1998: 89 Supplement 1,pp. S14-9). The model in Figure 2 views child development as an outcome of influences on health and, therefore, as being a characteristic of the population's health (e.g., a sub-category of functional status). The model attempts to include all of the important influences on child development (e.g., nutrition, early childhood experience and education, and social support).

include systems for providing clean drinking water, waste disposal, or vector control and community health promotion and education programs, including those designed to influence personal health practices, such as smoking, sexual practices, and physical exercise.

Collective lifestyles and health practices include diet, wellness behaviors, physical activity, sexual practices and abstinence, smoking, violent behavior, and substance use.[11,12] Collective lifestyles and health practices vary across communities as well as among individual members of those communities, and all of these attributes impact directly upon the population's health.

Model Overview – The Population's Health

The health of a population, the central focus of our model, is assessed using measures of disease, functional status, and well-being. These reflect both the level and distribution of health in a population. Measures of disease include rates of particular diseases within a population and how these rates vary over time, by place, and within subgroups of the population. The International Classification of Diseases provides a framework for undertaking and presenting these measurements.[13] Functional status refers to the ability of people to engage in the activities of daily living and social life.[14] The International Classification of Functioning, Disability and Health[15] presents a framework for assessing function that takes into account the social aspects of functional status and provides a mechanism to document the impact of the social and physical environment on a person's functioning. The concept of well-being reflects the subjective aspects of health. Unlike measures of disease and function, the measurement of well-being is less well defined and lacks an internationally agreed upon framework.

Mapping the current enterprise using the model

Table 1 provides for each component of the model an overview of typical approaches for gathering relevant information and associated data collection systems at national, state, and local levels. Table 1 indicates many strengths of the U.S. health statistics enterprise. The U.S. health statistics enterprise collects national data through internationally renowned ongoing national surveys, surveillance and registration systems, and administrative data systems. Some of these systems, and especially the surveillance, registration, and administrative data systems, also provide much needed state- and local-

[11]Frohlich, K. and Potvin, L., "Collective lifestyles as the target for health promotion," *Can J Public Health,* 1999: 90,Supplement 1, pp. S11-S13.

[12]Frohlich, K., Corin, E., and Potvin, L., "A theoretical proposal for the relationship between context and disease," *Sociology of Health and Illness*, 2001:23,6, pp. 776-797.

[13]World Health Organization, *International Statistical Classification of Diseases and Related Health Problems,* Revision10th rev. ed. (revised by the National Center for Health Statistics), NCHS, Hyattsville, Maryland, October 1998.

[14]Starfield, B., "Basic concepts in population health and health care," *Journal of Epidemiology and Community Health*, 2001;55(7): 452-4.

[15]World Health Organization, *International Classification of Functioning, Disability, and Health, short version*, World Health Organization, Geneva, 2001.

level data. Drawing upon Table 1, the following sections highlight examples of the strengths of the current U.S. health statistics enterprise, organized by the components in the model.

Current Enterprise – Context

Natural environment: The U.S. health statistics enterprise can draw upon ongoing national, state, and local monitoring of the natural environment, especially through systems maintained by the National Oceanographic and Atmospheric Administration, the National Weather Service, the U.S. Geological Survey, and the U.S. Environmental Protection Agency. These systems provide data on such influences on health as climate, topography, and air and water quality.

Cultural and political contexts: At the national level, the General Social Survey,[16] a personal interview survey of U.S. households conducted since 1972 by the National Opinion Research Center, provides data on various cultural, political, and economic aspects of society, such as political viewpoints, sociopolitical participation, and social trends.

Current Enterprise – Community Attributes

Health services: The U.S. has a wealth of high quality data collection systems providing ongoing national and state data on the structure, processes, and use of health services. These include, among others, the National Health Care Survey, the Annual Survey of Hospitals, the Healthcare Cost and Utilization Project, the National Medical Expenditure Survey, and the National Immunization Survey, which collect data through population, provider, client, and institutional surveys and ongoing administrative data systems. These systems provide data on aspects of health services, such as data on the adoption of new medical technologies and procedures, on the use of prescription drugs, and on the relationships among health insurance, access to care, health status, and health care costs and financing.

Built environment: Relevant ongoing data are available from Federal and state agencies. Especially noteworthy are ongoing data on land use, housing, air quality, and urbanization from the U.S. Department of Housing and Urban Development, the Census Bureau, and the U.S. Environmental Protection Agency.

Biological characteristics: Since 1790, the U.S. has collected decennial census data describing the age and sex distribution of the population at all geopolitical levels. Since the early 20th century, states have collected and provided to the Federal government data on births and deaths; again, these data are available at all geopolitical levels. The National Health and Nutrition Examination Survey, conducted by NCHS, collects detailed biological data pertaining to the health status of Americans on an ongoing basis.

[16]The National Opinion Research Center, *General Social Survey Codebook*, Available from: http://www.icpsr.umich.edu/GSS/ [accessed 1 September 2002].

Economic attributes: Ongoing and periodic data on employment, individual and household income, income inequality, and educational attainment are collected by the decennial Census of Population, the ongoing American Community Survey, and the Current Population Survey.

Population-based health programs: Studies by the National Association of County and City Health Officials, the Association of State and Territorial Health Officials, CDC, and others periodically assess the status of the public health infrastructure and the preparedness of state and local public health organizations.

Collective lifestyles and health practices: Ongoing data on health practices are collected by such national surveys as the National Health Interview Survey, while at the state level, the Behavioral Risk Factor Surveillance System and the Youth Risk Behavior Survey collect health practices data in each state.

Current Enterprise – The Population's Health

Disease: Disease incidence and prevalence data are collected through registries and reportable disease surveillance systems, population-based and event-based surveys, and administrative data systems. The Vital Statistics Cooperative Program provides nearly complete data at all geopolitical levels on mortality by cause. Morbidity data are collected through various national disease-specific surveillance systems with differing levels of completeness of case ascertainment, such as the HIV/AIDS Surveillance System and the National Notifiable Disease Surveillance System; these systems also collect data at all geopolitical levels. Other data on morbidity are collected through ongoing national surveys such as the National Health Interview Survey. Morbidity manifested in health care encounters is collected at the national level through surveys such as the National Ambulatory Medical Care Survey and the National Hospital Discharge Survey, and at all geopolitical levels through administrative data systems such as the Medicaid and Medicare files.

Functional status: At the national level, functional status data are collected through the National Health Interview Survey, the National Survey of Children with Special Health Care Needs, and systems designed to assess patients in nursing and home health care.

Using the model to describe issues and gaps

The examples of data collection systems provided in Table 1 indicate substantial unevenness in the capacity of the U.S. health statistics enterprise to provide data to characterize the health of its population and those factors that influence its health. While excellent and comprehensive data exist for some measures of the influences on health at some geopolitical levels, little or no ongoing data are available for other measures or

influences. An overview of the issues related to the current capacity of the U.S. health statistics enterprise to provide the information needed to develop a full understanding of its population's health is presented below. This overview is followed by a summary of the gaps in the capacity of the enterprise to provide the needed information. The reflections on issues and gaps are based upon the testimony delivered at our regional public hearings; local and national discussion groups; commissioned papers; the National Academy of Sciences/Committee on National Statistics workshop; and a review of Table 1.

Underlying issues

Analysis of the capacity of the U.S. health statistics enterprise to monitor the health of the U.S. population reveals several underlying issues that affect the health statistics for all aspects of health and its influences.

Broad definition of the population's health: The first issue is the failure of the health statistics enterprise to incorporate a broad definition of the population's health.[17] The multifactorial nature of the influences on the population's health is not currently reflected in our ongoing data collection systems, especially at the state and local levels.[18] Ongoing data collection systems are now largely disease-driven and disease-specific and generally do not operationalize health in a way that integrates disease, functional status, and well-being. Minimal ongoing data on functional status are collected, and almost no data are collected on well-being. Data sources that do exist do not lend themselves to linking and merging in a way that can easily address this multifactorial nature of health.

Units of analysis: Ongoing analyses conducted by state and national agencies focus largely on individual health and the determinants of individual health, rather than on populations and communities as units of analysis. Many state-level analyses reported in ongoing Federal publications focus on comparing individual health at the state level, rather than incorporating a population perspective to understand health.

State and local data: The availability of health statistics data varies greatly across geopolitical levels. During each of the local discussion groups and at all of the regional public hearings, repeated pleas were voiced by state and local public health practitioners, health care providers, and community organizations for more extensive state and local data, especially survey data.

Subpopulation data: Another theme repeatedly heard at the local discussion groups and regional public hearings revolved around the inconsistent quality and quantity of data on vulnerable populations.[19] This issue is particularly evident for subpopulations

[17]"Vision" discussion group, Albuquerque, 1 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/Albuquerque.pdf [accessed 1 September 2002]; "Vision" discussion group, Harrisburg, 11 November 1999, Available from: http://www.ncvhs.hhs.gov/hsvision/111699.pdf [accessed 1 September 2002].

[18]Fielding, J., National Center on Vital and Health Statistics public hearing testimony on: "21st Century Vision for Health Statistics," San Francisco, 30 October 2000, Available from: http://www.ncvhs.hhs.gov/001030tr.htm [accessed 1 September 2002].

[19]"Vision" discussion group, Albuquerque, 1 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision Albuquerque.pdf [accessed 1 September 2002]; Giachello, A., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Rosemont, Illinois, 10 July 2000, Available from: http://www.ncvhs.hhs.gov/000710tr.htm [accessed 1 September 2002].

that may be small in numbers nationally but relatively large locally as well as for subpopulations that may be relatively small in numbers in any individual local area but relatively large nationally. Also evident were concerns about the dearth of data on especially vulnerable subgroups within racial and ethnic minority populations, such as the minority elderly. Finally, concerns were expressed about inconsistent efforts to recognize and provide information on the diversity within minority populations, such as distinctions between subgroups of Asian Americans or Latino Americans. For example, while the new standard certificate of live birth promotes self-identification for Hispanics and Hispanic subgroups, no such self-identification of ethnicity is promoted for Blacks; similarly, many other data collection systems go no further than collecting broad data on Hispanic ethnicity but no data on specific Hispanic origin or ancestry.

Longitudinal and life cycle data: Partially due to not employing a broad definition of health, the U.S. health statistics enterprise does not adequately collect data describing the course of health throughout the lifecycle, and differential influences on health at different stages in the lifecycle.[20] When data collection systems exist in unconnected "silos," it becomes difficult or impossible to link those silos at the individual person level in order to develop over-time lifecycle health biographies. Different surveillance, administrative data, and survey systems can be difficult or impossible to link with each other for individual data subjects, even if the surveillance systems contain data on the same individual. Linkage of individual record data from different lifecycle stages at the person level across surveillance, administrative, and surveys is rare. Consequently, diseases and health events are largely analyzed as isolated episodes of illness.

Data availability versus data use: The U.S. health statistics enterprise is lacking both in the availability of data on some components of and influences on the population's health, and in the use of available data. Table 1 suggests the existence of a plethora of potentially important data sources available to the health statistics enterprise that are rarely utilized. This is especially true in regard to the lack of use of areal data pertaining to the natural and built environments.

Data gaps

This section provides an overview of the major gaps in available data for specific influences on and measures of the U.S. population's health.

Context: The U.S. health statistics enterprise collects and utilizes little data pertaining to either the cultural or political contexts of the population's health. Data collection systems provide minimal data on such important elements of the cultural context as racism, sexism, competition and cooperation, and norms and values. While data systems are used to measure the relationship of "race" and ethnicity to health outcomes, they generally do not include information needed to disentangle how

[20]Kalsbeek, W., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Research Triangle Park, 20 November 2000, Available from: http://www.ncvhs.hhs.gov/001120tr.htm [accessed 1 September 2002].

race/ethnicity, socioeconomic status, discrimination, race-as-perceived-by-others, and other factors affect health.[21] The enterprise generally fails to monitor the political context and rarely tracks on an ongoing basis the impacts of public policies and laws, the political culture, and political enfranchisement on the population's health.

Community attributes: Despite a substantial and increasing body of U.S. and international literature on the relationship of community social attributes to the health of individuals and populations, data on these attributes are not incorporated into ongoing analyses of health data at the national or state levels. The U.S. health statistics enterprise collects no ongoing data pertaining to such key community attributes as social cohesion, social influence, social networks, social support, and social change.

Similarly, data pertaining to economic attributes of communities are rarely incorporated into ongoing analyses of the health of their populations. While such ongoing surveys as the National Health Interview Survey and the Behavioral Risk Factor Surveillance System do collect data on individual and household income (although rarely on assets), disease- and condition-specific surveillance systems and administrative data systems do not collect such data.[22] Data on economic attributes are available on an ongoing basis from such data sources as the American Community Survey. Such data have been widely incorporated into research on the relationship of income inequality to a population's health, both in the United States and abroad. Yet ongoing analyses of the health of populations in the U.S. conducted by Federal and state governments do not regularly incorporate statistics on community economic resources.

The U.S. health statistics enterprise collects a wealth of data on the structure and processes of, and access to, health services. Yet these data are rarely related to the health of a population as the dependent variable. The contribution of our investments in health services to improvements in the population's health has not been thoroughly or systematically explored.

Data on the built environment are routinely available, but these data–pertaining to such key elements of the built environment as housing, school, transportation, communication infrastructure, and equity of access to these elements–are not used on an ongoing basis in analyses of health by national and state governments.

The relationship of population-based health programs to a population's health at the local or state level has also been largely absent from ongoing analyses.

The population's health: The operationalization and ongoing measurement of the U.S. population's health by the U.S. health statistics enterprise is largely confined to the incidence and prevalence of individual diseases and health conditions. Incidence and

[21]Krieger, N., "Discrimination and health," *Social Epidemiology,* Berkman, L. and Kawachi, I., ed. Oxford University Press, Oxford and New York, 2000.

[22]Krieger, N., Chen, J., and Ebel, G., "Can we monitor socioeconomic inequalities in health: a survey of U.S. health departments' data collection and reporting practices," *Public Health Reports*, 1997;112:481-491.

prevalence are generally measured in isolation, with little ability and few attempts to link incidence or prevalence of a specific disease to incidence or prevalence of another disease measured during the same life cycle stage through another data collection system, or to another disease or condition measured at a different life cycle stage. The measurement of functional status is confined largely to national surveys, and this important component of health is not typically addressed in ongoing surveillance or administrative data systems. Well-being is minimally measured in national surveys, and not measured at all in surveillance or administrative data systems. As a result, the working definition of health in the U.S. health statistics enterprise has been all but reduced to disease, which both reflects and reinforces the tendency of the health services system to focus on the treatment of disease.

The restricted working definition of health has important consequences for the practice of public health, leading to the concentration of public health programs on disease prevention and management, rather than on a more broadly integrated approach to promotion of health. This also limits the ability of decision-makers to employ a fully integrated model of the influences on the health of a population and its influences to help guide decision-makers. Making disease the primary focus of health services and public health programs oversimplifies the complex interactions that affect the health of people individually and the population as a whole.

Chapter Three: Developing, Evaluating and Improving Health Statistics

The health statistics enterprise and the Health Statistics Cycle

The health statistics needed to understand the dynamics of and influences on the U.S. population's health involve the efforts of a diverse and decentralized enterprise. Individually, the organizations and people that are involved in the American health statistics enterprise are responsible for specific data systems and focused efforts to use and disseminate data, with the measure of success often defined as meeting a specific need or use. The success of the overall health statistics enterprise (as defined in Chapter 1) depends upon the extent to which these individual components work in concert to meet the broader agenda of understanding the population's health and influences on it.

The performance of the health statistics enterprise, as well as of many individual programs that collect data and manage statistics, involves a number of key elements that are described in the "Health Statistics Cycle" (Figure 3). This model can be used to describe the steps required to define and meet data needs, to evaluate the strengths and weaknesses of current programs, and to suggest ways in which different components of the health statistics enterprise can be tied together into a more cohesive whole. The model draws upon the insights, suggestions, and observations of the broad range of participants in the health statistics vision process.

Health Statistics Cycle

Hub–Integrating the health statistics enterprise

Elements:

- Defining data needs
- Specifying necessary data attributes
- Identifying appropriate data sources and strategies
- Collecting, aggregating, and compiling data
- Analyzing statistics
- Translating statistics for users
- Evaluating extent to which needs are met

Elements in the Health Statistics Cycle

Figure 3 depicts seven elements in the health statistics cycle, arrayed in a circle, with each element leading to the next. Importantly, there is no first or last step; the model recognizes that the health statistics enterprise needs to be in a constant state of evolution, evaluation, and regeneration. The model is applicable both to the development of new statistical efforts and to efforts to make existing data systems that produce health statistics more effective. The model is also equally applicable to looking at the overall health statistics enterprise and to focusing on individual data systems or data needs.

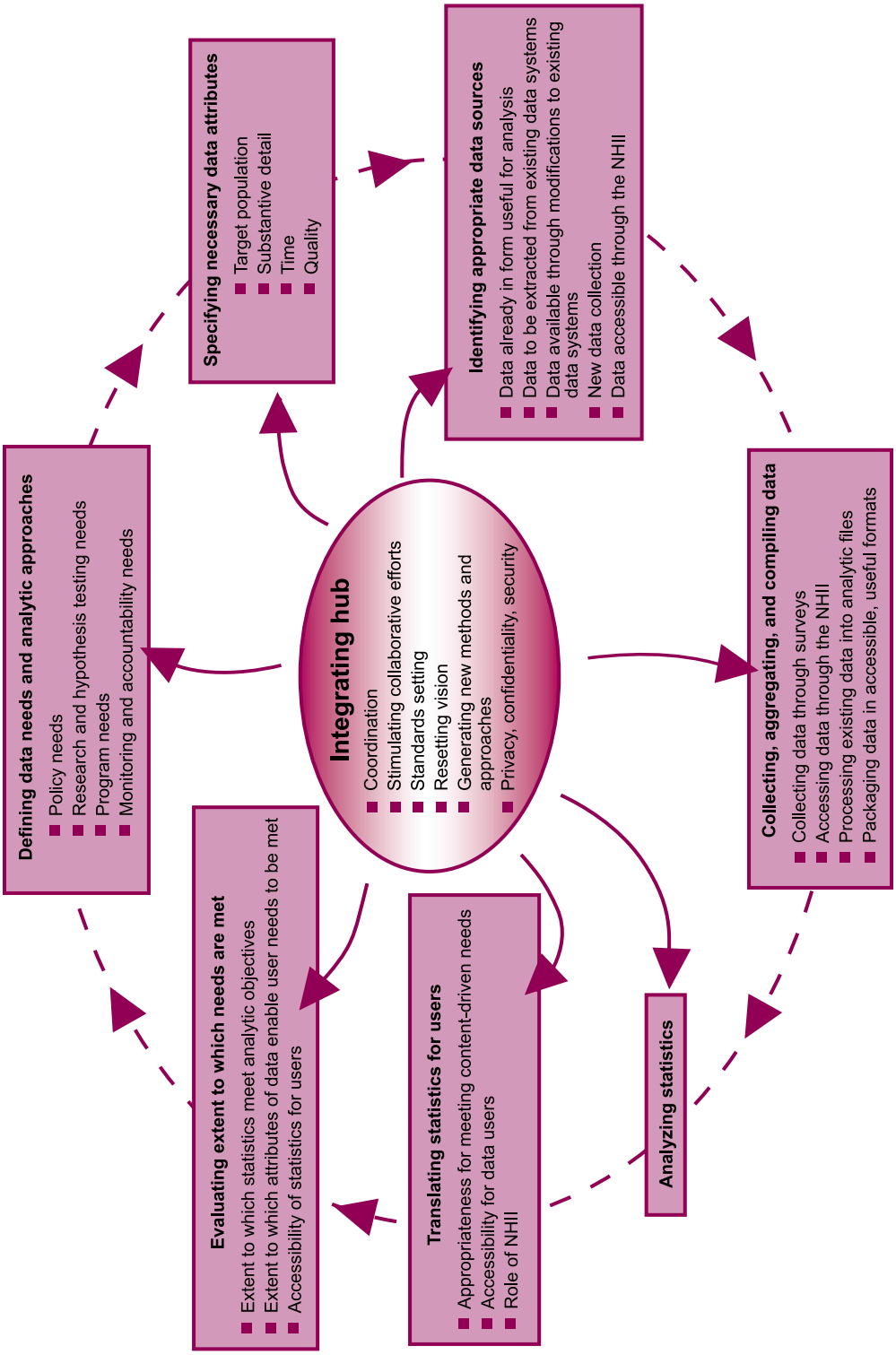


Figure 3. Health statistics cycle

Defining data needs

In the aggregate, the data needs to be met by the health statistics enterprise are determined by an ongoing and iterative process. The overall data needed to promote understanding of the population's health and the influences on the population's health are reflected by the scope of Figure 2. In working to address more specific needs, organizations that collect and compile statistics work with data providers (e.g., respondents to surveys, managers of health care records) and data users to focus on developing detailed definitions and specifications of data needs in areas including:

Health policies: for example, data to document the lack of health insurance or the relationship between the lack of health insurance and access to care;

Research and hypothesis testing: for example, information on the relationship between a risk factor and a health effect;

Health program needs: for example, data to define high-risk populations for an intervention;

Monitoring and accountability: for example, data that can be used in performance monitoring or tracking and reporting on high visibility health indicators; and

Health practitioners and health service providers: for example, data that can be used to assess the quality and efficiency of health care.

Specifying necessary data attributes

Each of these data needs can be further specified in terms of the data attributes required to meet the need. This greater specificity is needed so that alternative data sources or strategies can be evaluated, along with the process of matching data needs to available resources. Data attributes include the:

- Characteristics of the populations under study, such as racial/ethnic and socioeconomic subgroups, age/sex breakdowns;
- Degree of geographic detail, such as national, regional, community, or even more granular levels;
- Depth and complexity of information required to meet the need, such as number of questions to be asked regarding a particular topic in a telephone survey or the level to which data are coded in a classification system;
- Length of time to be observed, ranging from a single point in time to a full calendar year to a range of years;
- Timeliness of reporting of results, both in terms of the turnaround time for data processing and the currency of the time period for which data are collected;
- Type of respondent or data source, such as individual interviews, administrative records, or secondary use of data collected for another purpose;
- Measurement approach, such as direct observation, computer-assisted interview, telephone survey, laboratory analysis, and so forth;

- Extent to which observations need to be validated, verified, tested for reliability or subjected to other methods of ensuring data quality; and
- Degree of sensitivity and specificity needed to meet user needs.

Identifying appropriate data sources and strategies

With detailed specifications, various data sources or strategies can be evaluated in order to rapidly and efficiently meet user needs. Strategies, in increasing order of cost and complexity, may include:

- Identifying data that already exist in a format that meets user specifications and is already accessible for analysis;
- Extracting data from existing data sets;
- Identifying opportunities to use data initially collected for non-statistical purposes (such as administration or patient care);
- Adapting existing data collection systems to meet new needs; and
- Designing and implementing new data collection systems.

Collecting, aggregating, and compiling data

Depending on the strategy selected, this element in the cycle includes steps ranging from simple data acquisition to the management of complex, ongoing data systems:

- Establishing data use agreements with (or purchasing data from) organizations that manage existing data systems, such as an agreement between a state health department and HMOs in the state to compare the health of HMO enrollees and non-enrollees;
- Creating user-accessible analytic files from existing data sets;
- Aggregating statistics from multiple sources and creating new statistical products that, because of the synergy created by combining sources, have analytic usefulness beyond that of any one component data set;
- Managing ongoing surveys and other data collection systems to meet the needs of multiple users; and
- Launching surveys or other data collection approaches to meet specific user needs.

Analyzing statistics

Once data are in usable form, analysts must carefully define and apply analytic approaches. Individual components of the health statistics enterprise devote considerable resources to developing analytic packages and tools that can be directly applied to specific data sets, or data that are derived from individual data systems. Fewer efforts have been devoted to development of more widely applicable analytic tools and approaches that minimize the learning curve for individual users.

Translating statistics for use

Organizations that collect and compile health statistics bear responsibility for ensuring that the data collected are translated into forms that maximize their appropriate and successful use. Professional statisticians and data processing specialists typically staff organizations that collect and compile data, while many users are untrained in statistics and techniques for accessing them. The translation function seeks to bridge this gulf. For systems geared toward meeting narrowly defined needs of specific users, this element of the cycle involves a direct relationship between producer and user. More generally, the health statistics enterprise is responsible for steps including:

- Making data available in forms appropriate to the level of sophistication of potential users, ranging from raw micro-level data (for example, a detailed individual-level file from the National Medical Expenditure Survey that could be downloaded from the Internet for analysis and use in microsimulation modeling) to tabular or graphic presentations of the same data (for example, a table depicting health expenditures by source of payment);
- Providing easy access for potential users, for example through interactive World Wide Web systems designed to provide community users with access to health statistics data, such as Massachusetts' MassCHIP system, Missouri's MICA system, and Seattle/King County's VISTA/PH system.[23]
- Making it easier for users to locate available data, for example through World Wide Web portals that provide users with topical listings of data sources, such as www.fedstats.gov, a Federal government effort to direct users to statistical resources. In effect, this simplifies the task of "Identifying appropriate data sources" described earlier in the cycle;
- Providing technical assistance, training, and coaching to users, for example through distance training techniques such as World Wide Web-based interactive courses, through video-based training resources such as CDC's "Public Health Data: Our Silent Partner," through in-person training in conducting community health assessments, or through providing adequate explanations of statistical concepts such as confidence intervals in written reports.[24]

Many of these steps will be facilitated by further development of an overall National Health Information Infrastructure (NHII), which will provide greater interconnectivity between health providers and establish a framework for data access and exchange.[25]

[23]Friedman, D., Anderka, M., Krieger, J., Land, G., and Solet, D., "Accessing population health information through interactive systems: lessons learned and future directions," *Public Health Reports,* 2001;116:132-141.

[24]Centers for Disease Control and Prevention, *Public Health Data: Our Silent Partner*, Centers for Disease Control and Prevention, Atlanta, October 1999.

[25]National Committee on Vital and Health Statistics, *Information for Health: A Strategy for Building the National Health Information Infrastructure*, 2001, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

Evaluating the extent to which needs are met[26]

An integral component in the Health Statistics Cycle is the continuing evaluation of the extent to which needs are met and the identification of improvements that can be made. Without this explicit and self-conscious step, systems for collecting or compiling health statistics risk stagnation and irrelevance. This element involves multiple strategies and approaches, relying on direct feedback from users and extensive links outside the health statistics enterprise. Questions to be addressed in such an evaluation include:

- Are individual data collection systems meeting defined analytic objectives? (Are they addressing the questions asked?)
- Are the specific data attributes necessary or sufficient for addressing analytic objectives? (Are we collecting too much detail, or not enough?)
- Have analytic objectives changed, requiring change in (or elimination of) data collection systems? (If other systems have evolved since the initial selection of a data strategy, can they meet these needs?)
- Are data accessible enough for users, and are the tools available to support a level of analysis that meets user needs?
- Is the health statistics enterprise as a whole meeting its mission and addressing the broadly defined data needs described in Chapter One?
- Is the health statistics enterprise as a whole operating in an efficient manner?

Hub–integrating the health statistics enterprise

At the core of the health statistics enterprise is a set of activities and functions that collectively represent an integrating "hub." These efforts are key to ensuring that the diverse and decentralized set of organizations and individuals that are involved with health statistics can work in concert to meet a broader agenda. An effective hub, which may involve multiple organizations, approaches, and influences, can serve as a center of gravity for diverse elements, holding the health statistics enterprise together and ensuring that the cycle of development and improvement continues to move in constructive ways. If the movement from element to element in Figure 3 can be viewed as driven by the winds of a hurricane, the integrating hub is the "eye" around which these elements move.

Critical functions of the integrating hub include:

- Fostering information sharing among diverse players spanning different levels of government, public and private entities, etc.;
- Stimulating collaborative statistical efforts;
- Facilitating the process of locating and accessing data from diverse sources;
- Developing and maintaining common definitions, technical standards, classifications, methodological benchmarks, etc., that facilitate data sharing, systems integration, and comparability of data;

[26]Lafata, J., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Rosemont, Illinois, 10 July 2000, Available from: http://www.ncvhs.hhs.gov/000710tr.htm [accessed 1 September 2002].

- Renewing and revisiting the overall vision for the health statistics enterprise at defined intervals and evaluating whether the overall enterprise is meeting its mission;
- Stimulating research and development directed toward new methods and approaches;
- Providing guidance and policy on privacy, confidentiality, and security as they apply to health statistics and the organizations that collect, compile, and disseminate them; and
- Ensuring user and wider public input into decisions that affect the overall performance of the health statistics enterprise and of its individual component systems.

Mapping the current system using the model of the Health Statistics Cycle

The diversity and number of organizations and entities that are involved in the health statistics enterprise make a simple categorization difficult. A large number of organizations are involved in defining data needs, analyzing data, and using health statistics. A comparatively smaller number are involved in the direct operation of data collection systems, while still fewer are involved in the process of defining and integrating health statistics from an enterprise-wide perspective. In this section, these organizations are described in order to aid in identifying functional gaps in the overall health statistics enterprise.

Table 2 reflects the major functional elements in the Health Statistics Cycle, and provides examples of the organizations and entities carrying out these functions at the national, state, and local levels.

At the national level, data needs are often expressed in broad terms by the Congress and the Executive Branch; by advocacy groups; by the business community; by the research community; and by managers of Federal health programs. These broadly stated needs are translated into specifications for individual data collection systems by a wide number of public and private entities that are engaged in data collection and dissemination. Federal statistical agencies are charged with the continuing operation of large-scale national data systems that are geared toward meeting multiple (as opposed to specific, categorical) data needs, but these agencies are augmented by many other research and programmatic entities. NCHS is the only Federal agency with a singular mission of providing health statistics, but NCHS is actually responsible for only about one-tenth of spending on health statistics by its parent department, DHHS.[27] Other centers, institutes, and offices within CDC also maintain major ongoing surveys and

[27]Office of Management and Business, *Statistical Programs of the United States Government, fiscal year 2002*, Washington, D.C., 2001.

data collection systems. Ongoing national health-related data systems are managed by virtually every other agency within DHHS, and administrative data from the Medicare program and other federally administered programs are often made available for research and statistical purposes. Data from non-health agencies, such as the Census Bureau (for example, data on population demographics, health insurance coverage) and the Bureau of Labor Statistics (for example, data on work-related injury and employee health benefits) are important for the study of health, and private data sources, often funded by foundations, provide data frequently used by policy analysts and researchers. Professional associations, such as the American Medical Association or the American Hospital Association, provide widely used information from surveys of their members. Similarly, private organizations, often supported by foundations, frequently conduct significant efforts, such as the Center for Studying Health Systems Change's Community Tracking Study and the Urban Institute's National Survey of America's Families.

At the state level, the complexity of interaction within the health statistics enterprise is similar. All states have a vital statistics function and conduct population surveys of behavioral risk factors, maintain ongoing surveillance systems for communicable diseases, and collect data on newly diagnosed cases of cancer. Many state health departments have well-developed health data centers that produce or integrate data from multiple sources. As providers or financers of care, states maintain systems from which statistical information can be derived, collecting data as byproducts of the Medicaid and the State Children's Health Insurance Program (SCHIP) and patient-level data from the provision of clinical care. As protectors of the public's health and as regulators of health care providers, state agencies also maintain systems that gather public health surveillance reports as well as data collected through the oversight and enforcement arms of state government (licensing, certification, insurance oversight, worker compensation enforcement, overseeing nursing home quality, etc.). At the state level, non-public sources also play an important role; for example, in many states, hospital associations aggregate data on hospital discharges and make them available for statistical analysis.

At the local and community level, statistical efforts are much more episodic and sporadic. The health departments in larger cities and counties–such as New York City, Boston, Seattle/King County, and Los Angeles County–have well-developed health statistics offices. For the most part, due to the expense associated with primary data collection, smaller local governments and community organizations are consumers of data produced by others.

Finally, a number of organizations are involved in efforts to integrate and coordinate efforts across the health statistics enterprise. At the national level, the NCVHS provides advice to the DHHS Secretary and serves as a bridge to the private sector on statistical issues, while the internal DHHS Data Council brings together the DHHS agencies

to coordinate statistical efforts. Individual agencies have funded data planning and integration efforts, such as CDC's National Electronic Disease Surveillance System. National standards development organizations, such as the American National Standards Institute and its subject-matter oriented subgroups, facilitate processes for arriving at data standards that transcend public and private organizations and functions, and professional associations such as the National Association for Public Health Statistics and Information Systems, the Association of Maternal and Child Health Programs, the Council of State and Territorial Epidemiologists, and the Association of State and Territorial Health Officers, undertake data planning initiatives that have impact in all states. The Public Health Data Standards Consortium is a coalition of more than 30 national and state organizations committed to promoting and developing common data standards for pubic health and human services research. Both nationally and in many states, some health statistics agencies influence the coordination and integration of data collection efforts, including NCHS and some state centers for health statistics. Finally, foundations often provide leadership in data planning and coordination; for example, the Robert Wood Johnson Foundation has funded the Information for State Health Policy program and the recently established State Health Access Data Assistance Center.

Major issues and gaps in health statistics development, evaluation, and improvement

Participants in the vision process provided a great deal of positive feedback on the importance of health statistics in their work, examples of improvements in the availability and usability of selected data sets, and evidence of steps underway to improve the quality and relevance of health statistics. Many commented on the completeness and usefulness of data systems in the U.S., and some expressed the opinion that sufficient data were already available for a large number of specific policy and programmatic purposes. Particular attention was focused on the gains made in making data more widely available through the Internet and other electronic media, and on the nascent efforts to develop data standards through the Health Insurance Portability and Accountability Act (HIPAA) and other mechanisms.

Despite the strengths of the health statistics enterprise, and despite the masses of data collected by various systems, the participants in the vision process also repeatedly pointed out major gaps in the availability of data and in the ability of the current components of the health statistics enterprise to fully address needs expressed by data users. In a sense, the testimony and comments submitted by participants indicated that the health statistics enterprise is rich in data but poor in information. These comments, discussed in more detail below, focused on the difficulty in identifying appropriate data resources and in translating raw data into information useful for action at appropriate

geographic levels. The major issues, and resulting gaps in the health statistics cycle, are described here and addressed separately in recommendations from NCVHS (see page 49).

1) Insufficient connections between data sources (the suppliers of data, such as individual respondents to surveys, hospitals holding patient records, and so forth), organizations that collect or aggregate data (such as organizations that conduct surveys or compile data from administrative systems), and data users.[28] The historical view of health statistics has been limited to the collection, compilation, analysis, and processing of data – a static view that lacks explicit connections to users on both the front end (defining data needs and specifying necessary attributes) and the back end (translating statistics and evaluating the extent to which needs are met).

In recent years, there has been a constructive trend toward more formal reviews of data needs associated with important health issues. More often than not, however, these reviews focus on a specific subject matter area (such as a GAO study on environmental health statistics), or the statistical aspects are but one of many topics addressed in the context of a more broadly focused study of a health issue (for example, a discussion of issues in racial and ethnic data was included in the recent Institute of Medicine report on disparities in health care.[29,30] While these efforts are valuable, they are not useful in evaluating priorities across different health data needs or in integrating the efforts of data collection mechanisms that address multiple topics. Few are formally established as ongoing mechanisms that address enterprise-wide issues more broadly.

2) Lack of geographic detail and other specificity.[31] As noted in Chapter Two, it is clear that some data are available on most influences on the population's health and on many aspects of the working of the health services system. While such data may exist at the national level for many research purposes, it is clear that these data are not sufficiently detailed to meet most research, policy, or programmatic purposes.

The relative lack of data on major health issues and trends at the state and community levels is a clear concern, and participants in the vision process made clear that this lack of data has hampered the effective pursuit of public health goals.

[28]This issue was raised frequently in discussion groups sponsored by the vision development process in Albuquerque, NM; New Orleans, LA; and Harrisburg PA. Summaries are available at http://www.ncvhs.hhs.gov/hsvision/Abuquerque.pdf, http://www.ncvhs.hhs.gov/hsvision/NewOrleans1.pdf, and http://www.ncvhs.hhs.gov/hsvision/111699.pdf See also Rice, D, "Health Statistics: Past, Present, and Future." National Research Council . *Toward a Health Statistics System for the 21st Century: Summary of a Workshop*. Committee on National Statistics, Edward B. Perrin, William D. Kalsbeek, and Terri M. Scanlon, Editors. Division of Behavioral and Social Sciences and Education. Washington, D.C: National Academy Press, 2001; Atkinson, D., NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Washington DC, Available from http://www.ncvhs.hhs.gov/010111tr.htm; and Bethell, C,. NCVHS public hearing testimony on "21st Century Vision for Health Statistics, San Francisco, CA, Available from http://www.ncvhs.hhs.gov/001030tr.htm [all websites accessed September 2002].

[29]U.S. General Accounting Office, "Long-Term Coordinated Strategy Needed to Measure Exposures in Humans," Report GAO/HEHS-00-80, May 2000.

[30]National Academy of Sciences, Institute of Medicine, Smedley, B., Stith, A., and Nelson, A., ed., *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*, National Academy Press, Washington, D.C., March 2002.

[31]See Giachello, A., and Zangri, A., Chicago, NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Chicago IL, Available from http://www.ncvhs.hhs.gov/000710tr.htm [accessed September 2002]

Similarly, numerous studies have documented the lack of sufficient detail on race/ethnicity, income, and other socio-demographic factors in existing health statistics, even in large national efforts to collect health-related data. The lack of such detail has hampered research into the underlying causes of differentials in health; limited efforts to target interventions to groups at greatest risk for poor health; and made more difficult the task of monitoring and evaluating the effectiveness of public programs to improve health.

3) Lack of timeliness in making data available.[32] Data users express frustration with the chronic inability of existing systems at all levels of geography and detail to produce results that can be used to rapidly identify and address current health problems. While progress has been made in many areas (for example, NCHS has cut times to produce results in vital statistics and surveys; in the case of the National Health and Nutrition Examination Survey the time it takes to produce results was cut from 3 years to 6 months after a completed survey year), many systems lag decades behind in implementing technology that is in widespread use in other sectors. For example, state vital statistics systems often rely on either paper-based records or DOS-based PC programs to process some of the most widely used health statistics.

4) Existing data are unnecessarily difficult to locate, access, and use.[33] For the health statistics enterprise to be efficient and to avoid unnecessary duplication of effort, it is important to make maximum use of all relevant data that are collected. Three major impediments exist: First, knowledge of data resources is limited, and there are few comprehensive and readily available "guides" to finding existing data. While some data collection agencies make data available on the Internet in searchable formats, others do not; even when data are available, it is not readily apparent to many users how to search the Internet to find it. Data collected for specific analytic purposes–or for proprietary purposes–are often not widely publicized or referenced in searchable form. Second, once found, data are often not available in a user-friendly format, placing burdens on the user that can discourage use or make it more likely that a new data collection mechanism will be developed to meet the user's need. Making data widely and easily available can be expensive for data collection organizations, limiting access even to data collected for public purposes and for the public good. Third, barriers due to legal, administrative, and proprietary concerns as well as to organizational practices inhibit the use of existing data resources.

[32]See Lafata, J., NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Chicago, IL, Available from http://www.ncvhs.hhs.gov/000710tr.htm [accessed September 2002]

[33]"Vision" discussion groups, New Orleans, LA, 22 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/NewOrleans1.pdf, and Albuquerque, NM, 1 March 2002, available from: http://www.ncvhs.hhs.gov/hsvision/Albuquerque.pdf; Upshaw, V., NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Raleigh, NC, Available from http://www.ncvhs.hhs.gov/001120tr.htm [all websites accessed September 2002]; and Spar, E., written comments on Interim Vision Report.

5) Resource constraints have placed the performance and usefulness of major data collection systems at risk.[34] Multi-purpose, shared data systems maintained by public agencies are efficient mechanisms for meeting data needs. By sharing a basic infrastructure, these systems address multiple issues and can often be used to meet specialized needs for more detailed data. For example, many of DHHS's major surveys include a core of content designed to meet general needs, along with supplements or specialized components developed to meet a defined research or programmatic need. Yet these shared systems face chronic resource limitations, and have in some cases deteriorated to the point where they are at risk. Losing these workhorse data systems would cause an interruption in the data needed to monitor trends, lead to a lack of comparability across years, reduce opportunities for integration of diverse data activities, and spawn multiple, less efficient efforts more narrowly focused on specific needs or uses.

Reflecting these major issues, the vision process arrived at a consensus that existing enterprise-wide coordinating and integrating efforts were largely ineffective,[35] hampering potential collective efforts of the diverse elements in the health statistics enterprise to move toward needed improvements in health statistics. A great deal of effort has been devoted to such mechanisms in recent years, and organizations at the national and state levels have made progress in bringing about greater cohesion in the health statistics enterprise. Nonetheless, the need for new approaches was a clear theme throughout the vision process. A consensus exists that:

First, there is insufficient attention to developing consensus approaches and standards that would simplify collecting, protecting, and accessing data and make more efficient the process of developing methods, software, and specifications for individual data systems. Important topics to be addressed include, for example, standards for systems architecture, comparable questionnaires, common classification and coding systems, standards for the way in which data will be exchanged over the Internet, and conventions for presenting and formatting data that reduce the learning curve for users of multiple data sets.

Second, the health statistics enterprise, because it is decentralized as well as underfunded, has insufficiently invested in new methods (such as more efficient sampling approaches), technologies (such as Web-enabled, meta-data repositories, electronic data

[34]See Rice, D, "Health Statistics: Past, Present, and Future." National Research Council. *Toward a Health Statistics System for the 21st Century: Summary of a Workshop*. Committee on National Statistics, Edward B. Perrin, William D. Kalsbeek, and Terri M. Scanlon, Editors. Division of Behavioral and Social Sciences and Education. Washington, D.C: National Academy Press, 2001; and Lee, P., NCVHS public hearing testimony on "21st Century Vision for Health Statistics, San Francisco, CA, Available from http://www.ncvhs.hhs.gov/001030tr.htm [accessed September 2002].

[35]This observation based on extensive discussions in "Vision" discussion groups, notably Albuquerque, NM, 1 March 2002, available from: http://www.ncvhs.hhs.gov/hsvision/Albuquerque/pdf; Nelson, M., Lafata, J., and Merryweather, P, NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Chicago, IL, Available from http://www.ncvhs.hhs.gov/000710tr.htm; Hajat, A, NCVHS public hearing testimony on "21st Century Vision for Health Statistics, Washington, DC, available from http://www.ncvhs.hhs.gov/010111tr.htm; and Rodrian, M., NCVHS public hearing testimony on "21st Century Vision for Health Statistics, San Francisco, CA, Available from http://www.ncvhs.hhs.gov/001030tr.htm [all websites accessed September 2002].

access initiatives) and integrated approaches (e.g., combining survey field operations, unifying data entry systems for clients served by multiple state programs) that could reduce cost, improve data quality and availability, and provide added confidentiality protection.

Third, insufficient attention has been given within the health statistics enterprise specifically and within public health more generally to developing and implementing training in health statistics. For example, testifiers at the NCVHS public hearings could cite no U.S. school of public health that currently offers even an introductory course devoted to health statistics. Graduate level training is needed to introduce health statistics to masters and doctorate level students. Additionally, expanded in-service training is needed for professionals already engaged in any elements of the health statistics cycle, beyond the valuable training already offered to state department of health employees by NCHS. In order to be most useful, such in-service training should rely upon a range of training modalities, including short courses at regular professional meetings and Web-based courses.

Fourth, participants in the health statistics enterprise have not been effective at working with decision-makers who influence the processes that ultimately determine the availability of much of the data needed for studying health related issues. The sources for health statistics are often the by-product of decisions made primarily for non-health or non-statistical reasons: the structure or content of a medical claim; the laws and regulations governing access to medical records for research and statistical purposes; the standards for legal or administrative records; or the data elements included in data systems focused on the labor force or population demographics. Given the wide scope of influences on health, the ability to use non-health data sources is a strength of the health statistics enterprise; however, this often requires a major task in piecing together a patchwork of data resources. The current health statistics enterprise lacks the ability to develop and articulate effective positions and to engage with the producers of non-health sources of data that is important to understanding health, and also lacks the ability to effectively pursue opportunities to use data that flow from these other producers.

Fifth, the health statistics enterprise lacks an overall confidentiality and data-sharing framework that would facilitate the integration and linkage of information while protecting the privacy of individual data. Current systems for health data, as well as underlying health records, are covered by a patchwork of state and Federal laws and regulations. These protections rarely address issues associated with the creation and use of large databases with both the potential for important public purposes and the risk of individual disclosure.

Sixth, there is insufficient support, training, and assistance available to users of health statistics. Most data are released in a form that requires users to have a high level of statistical sophistication as well as direct experience with the individual data set.

Comparatively few resources are devoted to "help desks," to the development of materials accompanying a data file that give users a clear description of how the data were collected and its limitations, or to training of users. Many system managers invest resources in meeting the specialized needs of their own users, but few address more general issues of building a competent user community that can effectively use data resources to inform policy and program management. While NCHS offers the University Visitation Program and the biennial NCHS Data Users Conference, these programs are directed toward a relatively narrow band of technically sophisticated users.[36]

Seventh, a sub-optimal allocation of resources–or an undue burden on providers of data–results from the lack of coordination among elements of the health statistics enterprise. If users are unable to identify and use existing data or work with organizations to modify existing data systems, there is a tendency to initiate new, single purpose data collection systems. This can lead to duplication of effort, lost opportunities for efficient shared systems, and the failure to gain maximum value from data that are collected.

Eighth, the health statistics enterprise lacks flexibility and adaptability. For a number of reasons, it is difficult for organizations that operate systems that produce health statistics to quickly adapt to meet new needs. The funding for health statistics is distributed throughout the decentralized health statistics enterprise, and many key data sources are not principally designed to address health issues. As a result, it is often difficult to reallocate funding to emerging health needs. Managers of data systems must balance the need for stable trend data against the need to address new issues. Further, in the absence of new resources, data system managers may choose to tinker with existing systems rather than to introduce fundamental changes.

Finally, these issues–and the data gaps outlined in Chapter Two–are evidence of a fundamental lack of an enterprise-wide planning forum. Some entities have gained a measure of success in planning for specific elements of health statistics. For example, public/private standards development organizations are effective forums for developing consensus standards in specific areas, such as uniform billing records and standards for electronic transactions. The role of NCVHS has expanded in providing advice to DHHS on the implementation of the data standards aspects of recent Congressionally mandated administrative simplification initiatives. And the DHHS Data Council provides a forum for coordinating activities within the Department, playing a role in reviewing agency budget initiatives. However, there is no single place where these efforts are all tied together. While the existence of these and other separate planning and coordination efforts can be taken as an indication of the widespread view that such functions are needed, these multiple initiatives and forums themselves add to the perception of fragmentation and disorganization in the overall health statistics enterprise.

[36]National Center for Health Statistics, *Invitation and Registration*, Available from: http://www.cdc.gov/nchs/events/2002duc/invitation.htm [accessed 1 July 2002].

An effective planning and decision-making effort can improve integration and coordination, help set system-wide goals and expectations, establish common frameworks and standards, and find ways to develop funding sources to ensure that the enterprise meets the needs of users in an effective and efficient manner. NCVHS recommendations in the final section of this report address how this effort can be implemented.

Chapter Four: A Vision for the Health Statistics Enterprise

This chapter puts forth the joint vision of the NCVHS, CDC/NCHS, and the DHHS Data Council, described in terms of a mission, core values, and guiding principles. Achieving this vision will require the concerted efforts of multiple organizations and individuals at the national, state, and local levels, in both the public and private sectors. The partners in developing this vision look to the NCVHS for more specific recommendations on how to implement the vision, and to DHHS for continuing leadership and support from the national level for steps toward achieving the vision. NCVHS recommendations are included following this chapter.

Mission

The mission of the health statistics enterprise is to provide timely, accurate, and relevant information that can be used to improve the population's health, including information about the status of the population's health, information that can be used to formulate and evaluate the effects of health policy, and information that can be used to manage health interventions and programs. Health statistics must be easily accessible to a wide range of professional and community users.

Mission of the health statistics enterprise

To efficiently provide timely, accurate, and relevant information that can be used to improve the population's health.

The mission of the health statistics enterprise can be achieved through utilization of an overarching conceptual framework, adoption of four core values, and implementation of ten guiding principles.

Overarching conceptual framework

An overarching conceptual framework that helps to maintain a focus on needed data must guide the health statistics enterprise.[37] The conceptual framework should delineate major influences on health and place them in an overall context; emphasize the distribution and the level of health in populations; and yield a research agenda relevant to improving the health of populations. The use of an overarching conceptual framework will concentrate the health statistics enterprise on its mission, reduce efforts that are tangential to its mission, and help to accomplish needed changes within the enterprise. Most important, the use of

Attributes of an overarching conceptual framework

- Focuses on health, the population, and the community
- Emphasizes distribution and level of health
- Delineates major influences on health
- Defines a research agenda for improving the health of the population

[37] Black, C., Roos, N., Roos, L., "From health statistics to health information systems: a new path for the 21st century, Available from http://www.ncvhs.hhs.gov/hsvision/CP--Black.pdf [Accessed 18 October 2002].

such a framework will focus health policy-making on core influences on the health of populations.

Various alternative conceptual frameworks of the influences on the health of a population have been developed, especially in Canada and within the World Health Organization. Awareness of the historical and intellectual roots of these frameworks can help participants in the health statistics enterprise to better understand them, and why they include or omit various influences. A vigorous debate about the most appropriate conceptual framework and the redirection of focus that it might suggest could enliven the U.S. health statistics enterprise, as it has in Canada, Australia, and the United Kingdom and other European countries. Figure 2 may stimulate such a discussion.

Core values

The health statistics enterprise rests upon four core values:

1. Maintaining the confidentiality and security of individually identifiable health information
2. Maximizing the scientific integrity of all aspects of health statistics while acknowledging the specific ways in which the political, cultural, and business contexts may affect data collection, analysis, and interpretation
3. Optimizing the enterprise's accountability to its users to ensure the availability of the information that is needed for improving the nation's health
4. Ensuring the enterprise's accountability to its data suppliers to minimize their burden and to provide them with timely feedback.

Guiding principles

In order to actualize its four core values, the current U.S. health statistics enterprise must reinvigorate itself by implementing ten guiding principles:

1. Enterprise-wide planning and coordination to ensure relevance to local, state, and national policy and program decision-making and to an overall conceptual framework of the influences on the health of populations.

Enterprise-wide planning and coordination are needed to ensure that the individual components of the U.S. health statistics enterprise collectively fulfill the mission of efficiently developing, analyzing, and providing timely, accurate, and relevant information for improving the population's health. Without enterprise-wide planning and coordination, the individual components of the health statistics enterprise–those operating at the national, state, and local levels, and in both the public and private sectors–may undertake overlapping and duplicative activities. Enterprise-wide planning and coordination can help to concentrate resources on the highest priority policy and programmatic issues, can contribute to efficient use of scarce health statistics resources, and can maximize transfer of new methods and best practices.

Some planning and coordination already occurs, most notably at the national level. The CDC's National Electronic Disease Surveillance System activities, DHHS's Healthy People goal-setting efforts, and the implementation of HIPAA have engendered some public and private sector planning and coordination at all geopolitical levels. At the state and local levels, several attempts to foster planning and coordination have occurred, including Louisiana's Turning Point Partnership,[38] Arkansas' Health Improvement Process[39], and the National Association of County and City Health Officials and CDC's Mobilizing for Action through Planning and Partnerships process.[40] However, most planning and coordination efforts at all geographic levels are episodic and largely confined to compilation of data on individual diseases or groups of diseases. Current planning and coordination efforts tend to involve unequal dyadic relationships, with a Federal or state agency essentially determining the health statistics agenda, priorities, and methods of state or local health departments.

Enterprise-wide planning and coordination can help to concentrate resources on the highest priority policy and programmatic issues, can contribute to efficient use of scarce health statistics resources, and can maximize transfer of new methods and best practices.

Enterprise-wide planning and coordination will necessarily involve organizations and individuals at all levels of government and in the public and private sectors. Planning and coordination will entail establishing enterprise-wide expectations regarding standards for data collection, management, and exchange; for the timeliness, validity, and reliability of data; for access to and use of data; and for the ongoing evaluation of the individual components of the health statistics enterprise. Planning and coordination will also involve joint agenda-setting for health statistics organizations throughout the enterprise, in such areas as addressing new and emerging health problems, building sampling frames that can produce data generalizable to different units of analysis, expanding and validating the array of survey questions usable by different organizations within the health statistics enterprise, and integrating surveys, registries, surveillance systems, and administrative health data systems.

2. Broad collaboration among data users, producers, and suppliers at local, state, and national levels to ensure efficiency of the enterprise and the usefulness of the data that it produces.

The corollary to enterprise-wide planning and coordination is broad collaboration and integration to ensure efficiency of the enterprise's efforts and the usefulness of data

[38]Louisiana Turning Point Partnership, "Public health improvement plan: a catalyst for change," April 2000, pp. 38-40, Available from: http://www.lphi.org/turning.html [accessed 6 February 2002].

[39]Nugent, R., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Washington, D.C., 11 January 2000, Available from: http://www.ncvhs.hhs.gov/010111tr.htm [accessed 1 September 2002].

[40]Upshaw, V., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Research Triangle Park, 20 November 2000, Available from: http://www.ncvhs.hhs.gov/001120tr.htm [accessed 1 September 2002]; National Association of County and City Health Officials, Available from: http://www.naccho.org, [accessed 1 September 2002].

provided by the enterprise. Collaboration must occur in ways, among partners, and on issues that have not typically occurred in the past. Collaboration must occur at all geographic levels among those originally supplying data; those responsible for compiling, analyzing, and disseminating data; and those using the data, whether they are in the public or private sector.

Collaboration must rest upon mutual recognition of and respect for the needs of different organizations and different functions within the enterprise, as well as clear definitions of roles and responsibilities of organizations within the enterprise. Working toward the shared goal of producing information for improving the population's health, collaboration between those supplying data (e.g., community members and health care providers) and those who collect and process these data can result in increased awareness among data collectors of the demands and burdens on data suppliers and adjustment of data collection procedures to reduce those burdens. Collaboration among data users and data producers can also result in providing information that more closely meets the needs of data users.

3. Rigorous policies and procedures for protecting the privacy of individuals and the confidentiality and security of data.[41]

In protecting data confidentiality, we must distinguish between the confidentiality and security of individual record data and that of aggregate data. The most basic confidentiality and security issues around *potentially identifiable individual record data* include ensuring that the data are collected, released, and used only for specific approved purposes directly related to improving the population's health. The data must be held securely, so as to maximize both the privacy of the data subjects and the integrity of the data themselves. Only the minimum data necessary for specific approved uses must be employed, whether those uses involve creating knowledge, making health policy, or developing programs. The process for determining which uses of potentially identifiable individual record data improve the population's health, which uses are approved, and which data are minimally necessary may vary from data set to data set and from use to use. However, clear laws, regulations, or administrative or programmatic rules must delimit all uses of potentially identifiable individual record data.

The most basic confidentiality issues around the use of *aggregate data* include guaranteeing the privacy of individual data subjects through ensuring that their identities cannot be determined from electronic data files, tables, charts, maps, or written documents. This can be accomplished through a variety of disclosure limitation techniques, such as algorithms for suppression of cells with potentially high disclosure risk.

[41] "The terms *privacy, confidentiality,* and *security* are used in many different ways to discuss the protection of personal health information." In this report *privacy* refers to "an individual's desire to limit the disclosure of personal information." *Confidentiality* refers to "a condition in which information is shared or released in a controlled manner. Organizations develop confidentiality policies to codify their rules for controlling the release of personal information in an effort to protect patient privacy." *Security* consists of "a number of measures that organizations implement to protect information and systems. It includes efforts not only to maintain the confidentiality of information, but also to ensure the integrity and availability of that information and the information systems used to access it." Source: Committee On Maintaining Privacy And Security In Health Care Applications Of The National Information Infrastructure*, For the Record: Protecting Electronic Health Information*. National Academy Press, Washington, D.C., 1997, page 1.

The health statistics enterprise confronts new challenges in ensuring the confidentiality and security of individually identifiable health information. Some challenges will require novel solutions, and some challenges will require rededicating ourselves to stringent implementation of existing policies and procedures.

New confidentiality and security challenges include:

- the increasing availability of aggregated health statistics on the World Wide Web;
- the increasing interest in and opportunities for linkage of multiple data sets containing identified or de-identified individual records and for linkage of multiple data sets containing aggregated data;
- the increasing use of standard formats for data collection, enabling easier linkage of identified and de-identified individual record data;
- the emphasis upon greater efficiency in data collection through use of data for more than one purpose; and
- the need to harmonize health statistics confidentiality and security laws, policies, and practices in the private and public sectors at all levels of government with those mandated under HIPAA.

All of these challenges provide the health statistics enterprise with opportunities for more successfully fulfilling its fundamental mission. But addressing these challenges will require that the health statistics enterprise find and maintain an appropriate balance between–on the one hand–maximizing the confidentiality of data and the privacy of data subjects and–on the other hand–collecting and providing information for improving the population's health. Improved security systems can go a long way toward establishing confidence in the integrity of health statistics systems.

The health statistics enterprise must rely upon privacy, confidentiality, and security policies and procedures that are clear-cut, easily understandable to the public and to data subjects, and open to external review. Oversight of policies and procedures and their implementation is necessary, and external review of policies and procedures and their implementation should be encouraged or mandated.

4. Flexibility to identify and address emergent health issues and needs.

The health statistics enterprise must develop the capacity to identify health issues for the population both before and as they emerge. These issues could include the emergence of new at-risk subpopulations; new health services structures such as Integrated Delivery Networks,[42] expanded use of certain types of health services, such as alternative medicine approaches, and newly identified health needs. The enterprise should not restrict itself to continually refining retrospective measurements of already-identified health problems of the population. Ongoing data collection efforts must have the flexibility to prospectively identify new influences on the population's health, changing patterns in health, and health

[42]An "Integrated Delivery Network" is a network of affiliated health care providers and institutions with the capability of offering comprehensive and coordinated health services within a defined geographic area. Integrated delivery networks are often associated with a major metropolitan teaching hospital as the hub of the network and usually include physician offices, community health centers, and community hospitals. The network often has fully or partially centralized financial and administrative systems.

issues in new subpopulations. This flexibility should be incorporated into all ongoing data collection efforts, including registries and administrative data sets.

5. Use of data standards to facilitate sharing and comparability of data.

The recent NCVHS report *Information for Health: A Strategy for Building the National Health Information Infrastructure* describes standards as "the fundamental building blocks of effective health information systems…essential for efficient and effective public health and health care delivery systems."[43] In *Every Manager's Guide to Information Technology*, Peter Keen offers a more technical definition: "Standards are formats, procedures, and interfaces that ensure the equipment and software from a range of vendors will work together."[44]

Standards and consensus approaches can be developed for individual data elements (addressing, for example, definition, nomenclature, taxonomy, vocabulary, and coding); content of specific data sets; the structure of data files (field type and order, file formats, database structures); collection of data (questionnaire design, sampling protocols, laboratory methods, validation of data entry); storage, manipulation, and processing (security and integrity of data, rules for aggregating and analyzing); transmission of data (message format, format and content for each field, format for electronic data interchange security); and the presentation and dissemination of data.

Standards and consensus approaches for the 21st century health statistics enterprise must meet *five criteria*:

First, standards should be applicable to and implemented for multiple purposes on a truly enterprise-wide basis. The standards should apply to all geographic levels and to both the public and private sectors. The standards should cut across present disease- and issue- specific data silos, which have historically been fostered by categorical Federal funding streams. Standards should be equally appropriate for administrative health data systems, surveillance data systems, survey data, and registries.

Second, standards should extend beyond the health statistics enterprise itself to incorporate multiple purposes. In keeping with the recommendations in *Information for Health: A Strategy for Building the National Health Information Infrastructure,* standards should be equally applicable for the population health, personal health, and healthcare provider dimensions.[45]

Third, standards should facilitate data linkage and data sharing while protecting the confidentiality and security of identifiable individual records. Facilitation of data linkage should occur even in the absence of the implementation of the national health identifier

[43]National Committee on Vital and Health Statistics, *Information for Health: A Strategy for Building the National Health Information Infrastructure*, 2001, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

[44]Keen, P., *Every Manager's Guide to Information Technology*, Harvard Business School Press, Cambridge, MA, Available from: http://www.peterkeen.com/emgit002.htm [Accessed 18 October 2002].

[45]National Committee on Vital and Health Statistics, *Information for Health: A Strategy for Building the National Health Information Infrastructure*, 2001, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

initially mandated by HIPAA; data linkages can be accomplished at the community level even before they are accomplished at the individual level.

Fourth, standards should enhance the comparability of data and results across states, across governmental levels, and between the public and private sectors.

Fifth, standards should be flexible and responsive to the needs of different components of the health statistics enterprise, including local and state health needs.

In short, standards must be intelligently developed and applied for the 21st century health statistics enterprise. The development and application of these standards must constantly return to the touchstones: the purpose of the standards and the five necessary criteria.

The development of enterprise-wide standards and consensus approaches and the adoption of standards recommended by national or international organizations are facilitated by numerous factors. There is a rich and long history of the use of standards in health statistics, including the early 20th century development of the International Classification of Diseases and the U.S. standard certificates of live birth and death. Laws and policies, such as HIPAA and the Office of Management and Budget's release in 1997 of its revisions to the Standards for the Classification of Federal Data on Race and Ethnicity, mandate the use of standards for specific data collection and transmission efforts. Federal and state initiatives, such as the CDC's National Electronic Disease Surveillance System, its Public Health Conceptual Data Model and the activities of the Public Health Data Standards Consortium provide an impetus and a framework for the development of enterprise-wide standards. Finally, the continued recommendations, oversight, and encouragement of NCVHS itself will also prove essential to development and implementation of enterprise-wide standards and consensus approaches.

6. Sufficient detail at different levels of aggregation to support local, state, and national policy and programmatic decision-making.

Health policy-making and implementation, as well as the development and management of health programs, occur at the local, state, and national levels. Health statistics must support both policy and program functions at all three governmental levels. Similarly, health statistics must support research employing various units of analysis. Research involving health statistics can employ neighborhood, city or country, state, or national units of analysis. The specific policy, programmatic, and research uses of each data set must dictate the needed level of detail at different levels of aggregation.

The gold standard for data collection is a data set that can be "rolled up" seamlessly from the neighborhood level, to the city or county level to the state level to the regional level and to the national level. Such data sets allow the simultaneous development, implementation, and evaluation of policies and programs at the local, state, and national

levels. Some data collection systems of this kind currently exist within the U.S. health statistics enterprise, although they are almost exclusively limited to registries, surveillance systems, or administrative health data systems. Other existing data collection systems, such as population-based surveys, have the potential to be adapted to allow aggregation from local to state to national levels.

In developing data sets with sufficient detail at different levels of aggregation, consideration must be provided to both nationally small subpopulations with large local presences and subpopulations with small numbers in local areas but with measurable national presences. Examples of nationally small subpopulations with large local presences include many ethnic groups whose historical migration patterns have resulted in substantial concentrations in small numbers of U.S. cities or counties. Examples of subpopulations with small numbers in local areas but with measurable national presences include groups of individuals with specific chronic conditions or disabilities. Methodological issues inherent in developing data sets with sufficient detail to represent both of these types of subpopulations must be systematically addressed.

The ability to return useful information to data suppliers is an additional consideration in designing data sets at sufficient levels of detail. Through returning information to those who have employed their time, energy, and resources to supply data for statistical purposes, we can show the utility of health statistics generally as well as of a specific data collection effort. Demonstrating such a return on investment is a necessary step in improving the validity, reliability, and completeness of health statistics.

The U.S. health statistics enterprise must devote concentrated and coordinated attention to increasing the number of data sets providing detail at multiple levels of aggregation. This can occur only through the enterprise-wide planning, broad collaboration and integration, and compatible standards described in guiding principles #1, 2, and 5 above.

7. Integrated, streamlined data collection for multiple purposes.

Data collection within the current health statistics enterprise is distinguished by multiple, separate data collection systems. These systems operate at all governmental levels, cover many diseases, and use various types of data, including administrative health data, surveillance data, survey data, and data from disease or vital event registries. Because most of these systems are not integrated–or even coordinated–an individual data supplier often provides the same data to different data systems within the same health agency. For example, for each birth the obstetric hospital will provide overlapping data on the standard certificate of live birth, the hospital discharge record, claims forms for third party payment, newborn screening forms, the birth defects registry, and perhaps to condition-specific surveillance systems as well. All of these distinct data collection

systems require similar data on infant's and mother's name, place of birth, date of birth, address, race, ethnicity, and so forth, often including overlapping clinical data. Figure 4 uses a workplace injury as the health event to illustrate the wide range of health records that might be generated as an injured worker seeks treatment from an expanding range of health providers, along with administrative records related to the event, to payment for treatment, and to possible death. This demonstrates how multiple systems are used to aggregate health statistics. Few of these efforts to aggregate statistics can currently rely on standardized or integrated systems.

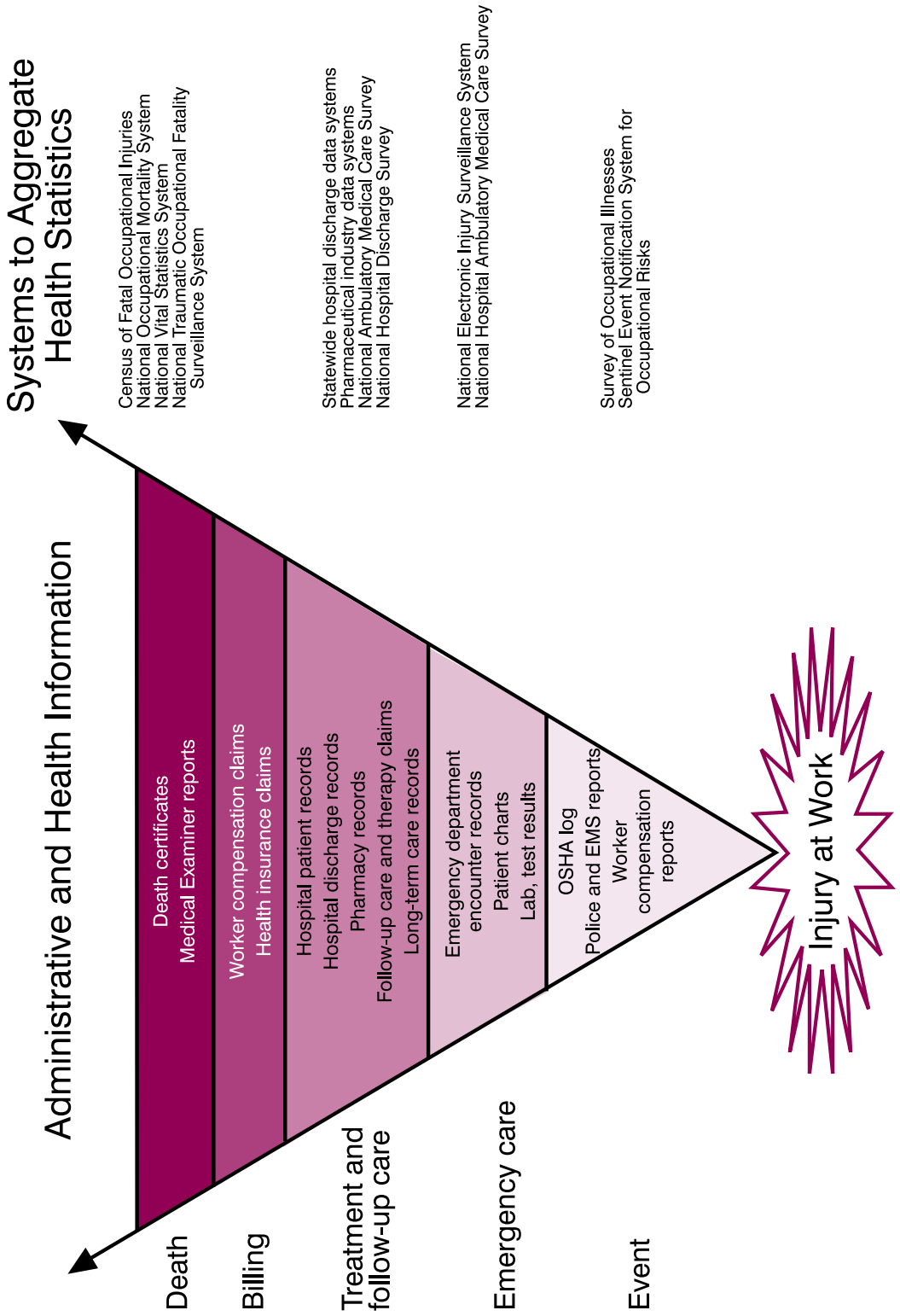


Figure 4. Multiplicity of data sources and statistical systems

The 21st century health statistics enterprise must rationalize and systematize presently disparate and distinct data collection systems. Multiple purposes or uses for data collection need not necessitate multiple data collection systems. By designing integrated, streamlined data collection for multiple purposes, greater efficiencies at reduced costs will be achieved.

The goals of integrated, streamlined data collection systems are threefold: First, to reduce the burden on data suppliers currently engendered by multiple overlapping data collection systems, by collecting the same data only once and then making it available to other data users;

Second, to minimize–where appropriate–the technical and administrative barriers between different data collection systems: administrative systems, surveillance systems, surveys, and registries; and

Third and finally, to enable the use of health data for different purposes by different users, rather than different uses requiring the establishment and maintenance of different data collection systems.

Integrated, streamlined data collection for multiple purposes need not imply a loss of confidentiality and security for individually identifiable data. With appropriate laws, policies, procedures, and technologies, the replacement of multiple data collection systems with integrated, streamlined data collection systems can enhance confidentiality and security.

8. Timely production of valid and reliable health statistics.

Timely production of valid and reliable data is inherent to the mission of the health statistics enterprise. Expectations and standards for the timeliness of health statistics have been changing over the last several years. Although in past years birth data were accepted as timely if published more than one year after the conclusion of the data year, individual states and NCHS now routinely publish birth data within 12 months of the conclusion of the data year. While previously cancer incidence data were routinely published two or even three years after the conclusion of the data year, CDC is now moving states toward publishing cancer incidence data within 12 months of the conclusion of the data year. The production of more timely health statistics may require careful consideration of whether the considerable investments currently made to achieve small increments in data quality are effective use of resources.

9. Appropriate access to and ease of use of health statistics.

To fulfill its mission, the health statistics enterprise must provide full and easy access to health statistics to a broad variety of users. The type and extent of the access must be appropriate for both the particular user and the particular use. Access to the detailed data needed by a health services researcher studying disparities across ethnic

groups in the utilization of different treatment modalities for advanced breast cancer would be inappropriate for community users conducting a local needs assessment.

The ready availability of data produced by the health statistics enterprise is a significant community resource for improving health. A wide variety of modes of access to health statistics must be employed, including the Internet, other forms of electronic access, public use data sets where feasible, and print media. Different modes of access must be viewed as complementary rather than as mutually exclusive. The increasing use of the World Wide Web as the backbone for interactive systems for disseminating public health data will provide the health statistics enterprise with new opportunities for customizing access to meet the distinct needs of different communities of users.[46] As is already the case in some of the more advanced Web-based interactive systems maintained by state health departments, customization would include providing alternative user interfaces for access to the same data, matched to different levels of user technical skills and goals.

10. Continuous evaluation of the scientific integrity, accuracy, and timeliness of health statistics and of the ability of the health statistics enterprise to fulfill its mission.

Continuous evaluation will aid in ensuring that new and emerging health statistics needs are identified and responded to on an ongoing basis. Evaluation efforts relevant to the health statistics enterprise are now largely limited to peer reviews of articles submitted to professional journals, the review process for competitive applications for Federal and foundation grants and cooperative agreements, and legislative review of health statistics programs. Mechanisms must be developed for continually evaluating the extent to which user needs are met, for assessing the scientific integrity, accuracy, and timeliness of health statistics, and for identifying improvements that must be made to meet future user needs. New metrics for evaluating the ability of the health statistics enterprise to meet future needs must be developed. New training is also needed so that those engaged throughout the health statistics enterprise are equipped with the appropriate methods to conduct continuous evaluation. Continuous evaluation must rest upon enterprise-wide planning and coordination in order to assure independence and openness.

As pointed out in Chapter One, the health statistics enterprise rests upon choices about what to study and how to study it. These choices are based on judgments about what issues merit data collection and analysis. Through aiming for and exercising self-awareness about values and choices, through inviting public scrutiny, and through following practices that ensure openness, the health statistics enterprise can maximize its scientific objectivity and its integrity.

[46]Friedman, D., Anderka, M., Krieger, J., Land, G., and Solet, D., *Accessing population health information through interactive systems: lessons learned and future directions*, Public Health Reports 2001;116: 132-141.

The health statistics enterprise and the NHII

In its recent report on *Information for Health: A Strategy for Building the National Health Information Infrastructure*, the NCVHS defines the NHII as including "not just technologies but, more importantly, values, practices, relationships, laws, standards, systems, and applications that support all facets of individual health, health care, and public health."[47] As emphasized in the NCVHS report, the purpose of the NHII is "sharing information and knowledge appropriately so it is available to people when they need it to make the best possible health decisions." The report points out that the NHII consists of three dimensions: the population health dimension, the healthcare provider dimension, and the personal health dimension. The health statistics enterprise draws upon data collected from all three NHII dimensions. The data utilized and analyzed by the enterprise fall within the loose boundaries of the population health dimension. As indicated in Figure 5, substantial overlap exists among the three NHII dimensions, and many health statistics fall in those areas of overlap. For example, vital statistics data can be used both for clinical purposes within the healthcare provider dimension and for health statistics purposes within the population health dimension. Mandatory disease reporting data can be used by authorized clinicians in the health care provider dimension, by affected individuals in the personal health dimension, and for health statistics within the population health dimension.

The 21st century health statistics enterprise in the U.S. must be viewed within the context of the NHII and its participants must become active and enthusiastic participants in the development of the NHII. Participation in the development of the NHII can help the health statistics enterprise to implement its core values and guiding principles. For example, the rigorous policies and procedures for protecting the privacy of individuals and the confidentiality and security of data developed within the health statistics enterprise can only be effective if they reflect those developed within the broader NHII. Similarly, the data standards utilized within the health statistics enterprise should wherever possible mirror those employed in the broader NHII. The NHII also provides the conceptual framework and ultimately the tools with which to transform integrated, streamlined data collection for multiple purposes from solely a guiding principle into a reality.

[47]National Committee on Vital and Health Statistics, *Information for Health: A Strategy for Building the National Health Information Infrastructure*, 15 November 2001, page 1, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

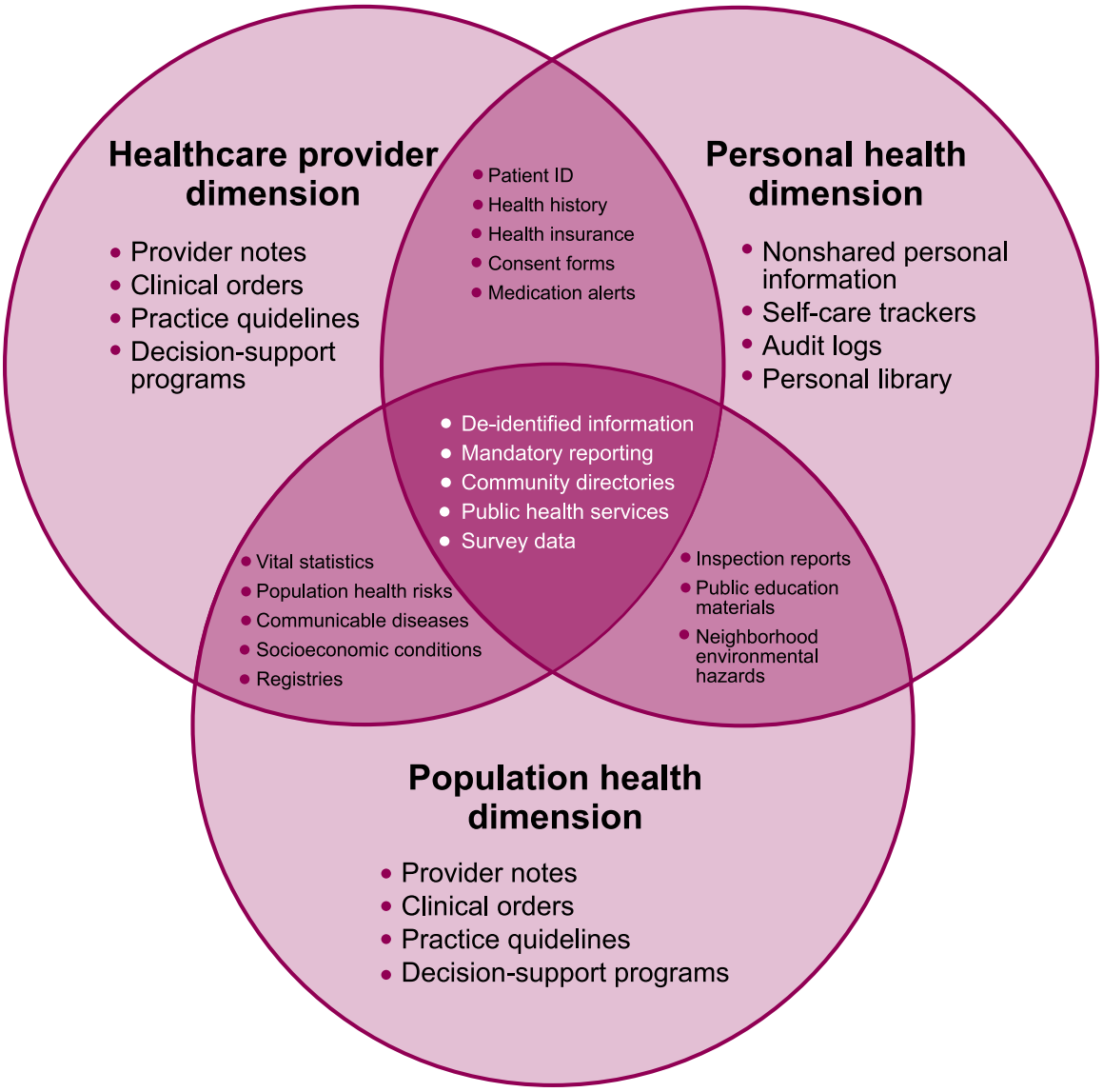


Figure 5. The National Health Information Infrastructure

NCVHS Recommendations for Achieving the Health Statistics Vision for the 21st Century

Developing the recommendations

NCVHS worked closely with its partners, CDC's NCHS and the DHHS Data Council, through a consensus process to identify a vision for health statistics. That shared vision is reflected in the preceding chapters of this report. Achieving this vision will require leadership from a variety of sources and resolution of a large number of scientific, interagency, and policy issues. As an advisory body charged with providing advice to the Secretary of Health and Human Services, NCVHS elected to exercise its independence in presenting its own recommendations for achieving this vision.

The NCVHS recommendations presented here are based on suggestions that emerged through the vision process. The recommendations address the problems identified in Chapter Two of this report relating to the lack of capacity of the U.S. health statistics enterprise to fully describe the population's health and the factors that influence the population's health. The recommendations also address the challenges outlined in Chapter Three relating to the ability of the enterprise to capture all appropriate data and to translate raw data into information useful for action.

Overview of recommendations

NCVHS recommendations are provided for implementing the vision described in Chapter Four. Thirty recommendations are included for moving the ten guiding principles from vision to reality, and six recommendations are included for moving the overarching conceptual framework for the health statistics enterprise from vision to reality. The recommendations relate to data access and use, data set development, data standards, enterprise structure, evaluation, policy, research agenda, and training.

NCVHS believes that in order to realize the vision priority must be placed on strategies for building a stronger integrating and coordinating "hub" for the health statistics enterprise, as described in Chapter Three. Implementation of the four NCVHS recommendations pertaining to enterprise-wide planning and coordination (Guiding Principle 1) is essential for constructing a strong health statistics enterprise "hub":

- Assignment of responsibility and authority for health statistics leadership within DHHS to a reconstituted National Center for Health Statistics, supported by a strong Board of Scientific Counselors. [GP1.1]
- Establishment of a national Health Statistics Planning Board with overall responsibility for planning and agenda setting for health statistics in the United States. [GP1.2]

- Designation of overall responsibility for health statistics planning and coordination activities within each state's government to a single state agency, supported by autonomous state Health Statistics Planning Boards. [GP 1.3]
- Development of appropriate graduate and in-service training and continuous education addressing all elements of the health statistics cycle. [GP 1.4]

These recommendations require structural and cultural changes within the health statistics enterprise, starting at the Federal level and extending both to state governments and the private sector. NCVHS regards it as critical for DHHS to assert leadership by quickly establishing a locus of responsibility for follow-up on these recommendations, addressing organizational issues raised in these recommendations, and providing for adequate advisory input and structures for collaboration across elements of the enterprise.

Recommendations relating to Guiding Principle 1 are essential first steps toward achieving the health statistics vision, and are highlighted because of their importance. However, they only enable the achievement of the nine other guiding principles and the overarching conceptual framework. The remainder of this chapter provides detailed NCVHS recommendations. Prior recommendations from the NCVHS and other sources that were developed during earlier efforts to assess population health information needs in the U.S. and other countries are referenced where appropriate.

Detailed recommendations for achieving the ten guiding principles

Guiding principle 1. Enterprise-wide planning and coordination to ensure relevance to local, state, and national policy and program decision-making and to an overall conceptual framework of the influences on the health of populations.

GP recommendation 1.1: Enterprise structure – reconstituted NCHS, with a strong Board of Scientific Counselors

Planning and coordination among Federal agencies contributing to the health statistics enterprise must be significantly strengthened to ensure that duplication and overlap are minimized and that the Federal government efficiently and effectively advances the mission of the health statistics enterprise.[48] Historically, this goal has been elusive due to the highly decentralized nature of Federal health statistics programs. Achieving this goal requires that the Secretary should designate a reconstituted NCHS as the lead agency for health statistics within DHHS, with responsibility and authority for statistical leadership, health statistics standards, health statistics coordination, and health statistics

[48]This point was made repeatedly by participants in the process for developing health statistics vision for the 21st century. For example, see: Rice, D, "Health Statistics: Past, Present, and Future." National Research Council. *Toward a Health Statistics System for the 21st Century: Summary of a Workshop*. Committee on National Statistics, Edward B. Perrin, William D. Kalsbeek, and Terri M. Scanlon, Editors. Division of Behavioral and Social Sciences and Education. Washington, D.C: National Academy Press, 2001; and comments by Atkinson, D., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Washington, D.C., 11 January 2000, Available from: http://www.ncvhs.hhs.gov/010111.tr.htm [accessed 1 September 2002].

agenda planning.[49] In order to maximize its contribution to the NHII generally and the health statistics enterprise specifically, the reconstituted NCHS should be strengthened so that it can provide greater national leadership for integrating and coordinating the health statistics enterprise, serving as an integral part of the "hub" described in Chapter Three. The responsibilities of the reconstituted NCHS must stretch far beyond the collection and analysis of a limited number of data sets and include the interpretation of data on population health and influences on population health from a wide variety of DHHS and non-DHHS Federal agencies, state agencies, and private organizations. The reconstituted NCHS must establish a strong ongoing link to any DHHS office and senior position instituted (as recommended previously by the NCVHS) to provide strategic leadership for developing and setting the agenda for the NHII. The reconstituted NCHS must also work closely with its counterpart agencies in the states, providing support and leadership to help states fulfill their potential as key components of the health statistics enterprise.

The reconstituted NCHS must be supported by a Board of Scientific Counselors, established by the Secretary of Health and Human Services.[50] This board would advise NCHS on carrying out its expanded responsibilities. In close coordination with NCVHS (which would continue to advise the Secretary on overall health information policy, HIPAA implementation, and related health statistics activities), the Board would provide direct advice and oversight for NCHS in the conduct of its statistical programs, in efforts to integrate statistical efforts across DHHS and the wider Federal government, and in constructive roles that NCHS and its partners can play in advancing the overall vision for the health statistics enterprise.

[49]Prior recommendations and related models:

a) NHII recommendation 1 for the federal government. Source: NCVHS, *Information for Health: A strategy for building the National Health Information Infrastructure*, U.S. Dept of Health and Human Services, Washington, D.C., November 2001, page 39, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf

b) NHII recommendation 2 for the Federal government. Source: NCVHS, *Information for Health: A strategy for building the National Health Information Infrastructure*, U.S. Dept of Health and Human Services, Washington, D.C., November 2001, page 40.

c) Recommendation on "strengthening the statistical system through organizational change." Source Norwood, J. *Organizing to Count: Change in the Federal Statistical System.* Urban Institute Press, Washington, D.C., 1995, pp. 71-76.

d) Bradburn, N. "The future of federal statistics in the information age," *The 1997 Morris Hansen Lecture*, 22 October 1997, page 14.

e) Recommendation for new Health Information, Communication, and Data Agency. Source: Milbank Memorial Fund, Bufford, J. and Lee, P., "Health policies for the 21st century: challenges and recommendations for the U.S. Department of Health and Human Services," September 2001, page 18. Available from: http://www.milbank.org/01091healthpolicies.html [accessed 7 November 2001].

[50]In 2000, a charter was approved for a Board of Scientific Counselors for NCHS; however, the Secretary has not yet appointed members. The charter and membership for this Board would need to be revisited in light of the additional responsibilities proposed for NCHS in these recommendations. For example, the expanded responsibilities of the reconstituted NCHS will necessitate that the Board's membership be drawn from a broad spectrum of the health statistics enterprise, including NCHS' state counterpart agencies.

GP recommendation 1.2: Enterprise structure – autonomous Health Statistics Planning Board

Planning and coordination between private and public sectors and Federal, state, and local levels of government should occur through a national Health Statistics Planning Board (HSPB).[51,52] The HSPB should be initially Federally funded but autonomous. Its membership should consist of representatives from throughout the health statistics enterprise, including representatives of organizations that serve as data sources; representatives of organizations that collect, aggregate, and compile data; representatives of Federal, state, and local governments; data analysts; and a broad range of health statistics users. Consideration should be given to a high level Federal mandate that would establish its roles and responsibilities, but the Board would ultimately be a collaborative effort of multiple member organizations. The HSPB will be an enterprise-wide forum that fosters intersectoral collaboration and integration. Its functions should include:

- establishing the agenda for the enterprise and identifying high priority population health issues requiring data and analysis;
- periodically evaluating the extent to which the enterprise is fulfilling its mission;
- resetting the vision as necessary;
- engaging in consensus planning for correcting deficiencies in the enterprise's capacity to fulfill its mission;
- recommending adequate resources to support the enterprise;
- proposing changes in existing data collection systems, new uses of existing data collection systems, and new data collection systems to provide needed information on high priority population health issues; and
- fostering the development of enterprise-wide standards.

[51]The NCVHS reports to the Secretary of DHHS and provides overall advice on health information policy. The BSC would report to the Director of NCHS, and provide advice on NCHS activities. The HSPB would be a non-governmental forum, providing overall direction to all levels of the health statistics enterprise in the public and private sectors. As the health statistics enterprise evolves and is strengthened, the respective roles of related advisory bodies will need to be assessed.

[52]Prior recommendations and related models:

a) National Research Council, Kalsbeek, W. and Perrin, E., "Some new directions," *Toward a Health Statistics System for the 21st Century: Summary of a Workshop*, National Academy Press, Washington, D.C. National Academy Press, 2001, pp. 62-64.

b) NHII recommendation 1and 2 for state and local government. Source: NCVHS, *Information for Health A strategy for building the National Health Information Infrastructure*, U.S. Dept of Health and Human Services, Washington, D.C., November 2001, page 42.

c) Australian Institute of Health and Welfare, "Developing public health information capacity, an infrastructure for managing the national development of public health information;" *National Public Health Information Development Plan*, recommendation 3.5, 1999; Source: Australian Institute of Health and Welfare and the National Public Health Information Working Group, *National Public Health Information Development Plan*, Australian Institute of Health and Welfare, Canberra, 1999, page vi.

d) National Health Information Agreement for Australia. Source: Australian Institute of Health and Welfare, "Sharing the fruits of coordination of health statistics in Australia," presented at Rome, Italy, 14-16 October 1998.

e) *A Network for Health Surveillance in Canada*. Source: Health Canada, *A Network for Health Surveillance in Canada*, July 1999.

f) Recommendation on Canadian Population Health Information structure. Source: Canadian Institute for Health Information, *Roadmap Initiative . . .Launching the Process*, Canadian Institute for Health Information, Ottawa, March 2000, pp. 19-21.

GP recommendation 1.3: Enterprise structure – coordination of state health statistics activities by a single state agency, supported by autonomous state Health Statistics Planning Boards

Within each state, planning and coordination among state agencies contributing to the health statistics enterprise should be assigned to a single state agency. This agency should have the leadership and technical expertise needed for state government-wide planning, budgeting, and agenda setting. Such agencies, often referred to as State Centers for Health Statistics, should be parallel in responsibility and capability to the strengthened Federal counterpart described in Recommendation GP 1.1. These agencies should be supported by autonomous state HSPBs, which will provide parallel functions at the state level to those described for the national HSPB in GP recommendation 1.2.[53]

GP recommendation 1.4: Training – expanded graduate, in-service, and continuous education training opportunities in health statistics

Expanded training opportunities should be developed and offered addressing all elements of the health statistics enterprise. The national HSPB, together with the reconstituted NCHS and its Board of Scientific Counselors, should assume leadership for conducting an enterprise-wide training needs assessment and for developing a training strategic plan.

Guiding Principle 2. Broad collaboration among data users, producers, and suppliers at local, state, and national levels to ensure the efficiency of the enterprise and the usefulness of the data that it produces.

Recommendations GP 1.1–1.3 address this principle.

[53]Prior recommendations and related models:
a) NHII recommendation 1 for state and local government. Source: NCVHS, *Information for Health: A strategy for building the National Health Information Infrastructure*, U.S. Dept of Health and Human Services, Washington, D.C., November 2001, page 42, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].
b) Recommendation for Health Assessment Panel. Source: Louisiana Turning Point Partnership, "Public health improvement plan: a catalyst for change," April 2000, pp. 38-40, Available from: http://www.lphi.org/turning.html [accessed 6 February 2002].
c) Center for Health and Environmental Statistics, Kansas Department of Health and the Environment, Phillips, L., "Health care data governing board," January 2000 presentation.
d) Kansas Health Care Data Governing Board, "Executive Summary," *Kansas Health Care Data Governing Board Annual Report 2001, 2001. Data Governing Board Annual Report, 2001.*

Guiding Principle 3. Rigorous policies and procedures for protecting the privacy of individuals and of the confidentiality and security of data.

GP recommendation 3.1: Policy – Code of Fair Information Practices

A model code of fair information practices for the health statistics enterprise should be debated and developed by the national HSPB, focusing on:

- individual rights of access to data pertaining to themselves;
- procedures for authorizing disclosure of individually identifiable data from data subjects and data sources;
- model privacy impact statements for new data collection systems, data linkages, and reuse of data collected for purposes other than those specifically mandated in statutes;
- policies related to the re-release of identifiable, de-identified, or anonymous data for policy, programmatic, or research purposes;
- re-release of identifiable, de-identified, and anonymous data for any purposes other than policy, program, and research;
- accountability for inappropriate use of identifiable and de-identified data;
- oversight of privacy, confidentiality, and security procedures and practices within the health statistics enterprise, including the roles of human rights review committees and privacy boards such as those approved by the Office of Human Research Protections; and
- public responsibility for determining a proper balance between use of identifiable and de-identified data, on the one hand, and societal priorities for improving the population's health through improving health statistics, on the other hand.[54]

GP recommendation 3.2: Policy - HIPAA guidance on privacy

Specific guidance should be developed by the Office for Civil Rights within DHHS to make clear how the Standards for Privacy of Individually Identifiable Health Information[55] adopted under the HIPAA relate to health statistics. This guidance should specify responsibilities of public and private producers of health statistics for protecting records and should provide clarity to health care providers on the appropriate uses of health information for statistical purposes under various elements of the Standards.

[54]Prior recommendations and related models:
a) Recommendations 1 and 3.1, source: National Research Council, *For the Record: Protecting Electronic Health Information*, National Academy Press, Washington, D.C., 1997, pp. 7-13.
b) Confidentiality Recommendations, source: National Committee on Vital and Health Statistics, "Health Privacy and Confidentiality Recommendations," 25 June 1997."
c) Bruening, P., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Washington, D.C., 11 January 2000, Available from: http://www.ncvhs.hhs.gov/010111tr.htm [accessed 1 September 2002].

[55]Department of Health and Human Services, "Standards for Privacy of Individually Identifiable Health Information; Final Rule," 45 CFR Parts 160 and 164, Federal Register: August 14, 2002, Volume 67, Number 157, pp. 53181-53273.

GP recommendation 3.3: Policy – model privacy, confidentiality, and security statutes

Model state and Federal privacy, confidentiality, and security statutes should be developed for public sector–sponsored health statistics efforts. The model statutes should apply to a wide range of data collection systems and uses of health statistics. The model statutes should provide definitions of key terms, establish protection from subpoena of data for statutorily mandated data uses and users, specify rules and procedures for re-release of identifiable data specifically collected for health statistics purposes, and specify penalties for misuse of data.[56]

GP recommendation 3.4: Policy – best practices privacy, confidentiality, and security

A toolbox of privacy, confidentiality, and security best practices should be developed for use throughout the health statistics enterprise, based upon the code of fair information practices, the DHHS guidance on the relationship of the HIPAA privacy standards, and the model statutes.[57]

GP recommendation 3.5: Training – privacy, confidentiality, and security

A model training agenda and training materials on privacy, confidentiality, and security for use throughout the health statistics enterprise should be developed to accompany the code of fair information practices and the tool box of privacy, confidentiality, and security best practices.[58]

GP recommendation 3.6: Research agenda – privacy, confidentiality, and security

A research agenda for the study of privacy-, confidentiality-, and security- enhancing statistical methods and technologies should be developed and Federally funded, addressing such issues as disclosure avoidance and cell size suppression, re-release of aggregated data through interactive Web systems, techniques for de-identifying or

[56]Prior recommendations and related models: Norwood, J., "an alternative approach if needed: passage of a new national statistical law," *Organizing to Count: Change in the Federal Statistical System*, Washington, D.C., 1995, pp. 76-80.

[57]Prior recommendations and related models: Klein, S., letter to NCVHS, 14 November 2001.

[58]Prior recommendations and related models:
a) Recommendations 3.4 and 5.2, source: National Research Council, *For the Record: Protecting Electronic Health Information*, National Academy Press, Washington, D.C., 1997, pp. 13-14, 16-17.
b) Kibbe, D., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Research Triangle Park, 20 November 2000, Available from: http://www.ncvhs.hhs.gov/001120tr.htm [accessed 1 September 2002].

making data anonymous, and ways of improving access to data for analytic purposes without sharing of identifiable information.[59]

Guiding Principle 4. Flexibility to identify and address emerging health issues and needs.

GP recommendation 4.1: Data set development – early warning systems

Early warning systems should be developed that actively monitor the population's health and potential influences on the population's health in order to identify emerging issues requiring data collection and analysis. These warning systems may rely on less statistically rigorous data than other data collection systems utilized by the health statistics enterprise.

GP recommendation 4.2: Data set development – flexible data collection

Organizations responsible for ongoing data collection systems that feed data into the health statistics enterprise should develop mechanisms to regularly and rapidly modify the specification of target populations, data collection strategies, and data collection instruments in response to emerging issues. Revision cycles for these ongoing data collection systems should be shortened so that modifications can occur more rapidly and more flexibly.

GP recommendation 4.3: Research agenda – emergent health needs

A research agenda should be developed and funded to:

- explore new data collection strategies that can rapidly and flexibly provide data on emerging influences on the population's health.
- assess the validity and reliability of items used in key ongoing data collection systems that feed the health statistics enterprise.[60]
- estimate any loss in accuracy from early publication of provisional, incomplete data from ongoing data collection systems.

[59]Prior recommendations and related models:
a) Recommendation 5-6, source: Institute of Medicine, *Protecting Data Privacy in Health Services Research*, National Academy Press, Washington, D.C., 2000, pp. 84-88.
b) Bruening, P., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Washington, D.C., 11 January 2000, Available from: http://www.ncvhs.hhs.gov/010111tr.htm [accessed 1 September 2002].

[60]Prior recommendations and related models: Fowler, J., letter to NCVHS, 13 November 2001.

Guiding Principle 5. Use of data standards to facilitate sharing and comparability of data.

GP recommendation 5.1: Data standards – harmonizing NEDSS and the overarching conceptual framework

The public health conceptual data model and the National Electronic Disease Surveillance System (NEDSS) logical data model developed by the CDC should be harmonized with the overarching conceptual framework for the health statistics enterprise in order to facilitate the development of standards.

GP recommendation 5.2: Data standards – population health

Enterprise-wide standards should be developed for classifications of the influences on population health, for data elements commonly used in all data collection modes (such as demographics and identifiers), for methods of electronically transmitting data, for presenting and disseminating data, and for providing electronic access to data. Standards should be common across all geographic levels. These standards should reflect where possible and be coordinated with those adopted for the NHII, and developed by private sector standards development organizations and international organizations.[61]

GP recommendation 5.3: Data Standards – implementation guidance

Implementation guidance and procedures and provisions for implementation audits should accompany all standards developed for use in the health statistics enterprise.

[61] Prior recommendations and related models:

a) NHII recommendation 4 for Federal government, source: NCVHS, *Information for Health: A strategy for building the National Health Information Infrastructure*, U.S. Dept of Health and Human Services, Washington, D.C., November 2001, page 41, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

b) NHII recommendation 2 for state and local government, ibid, page 42.

c) Recommendation on integration of health information systems, source: National Institute of Standards and Technology, *Highway to Health: Transforming U.S. Health Care in the Information Age–Executive Summary*, 1996, pp. 13-14.

d) Recommendation on common data standards, source: Canadian Institute for Health Information, *Health Information Roadmap: Responding to Needs*, Canadian Institute for Health Information, Ottawa, 1999, page 9.

e) Rodrian, M., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," San Francisco, 30 October 2000, Available from: http://www.ncvhs.hhs.gov/001030tr.htm [accessed 1 September 2002].

f) Klein, S. letter to NCVHS, 14 November 2001.

GP recommendation 5.4: Data standards – participation in standards development

Participants in all components of the health statistics enterprise should actively contribute to the national standards setting process, through such mechanisms as the Public Health Data Standards Consortium.[62,63]

Guiding Principle 6. Sufficient detail at different levels of aggregation to support local, state, and national policy and programmatic decision-making.

GP recommendation 6.1: Research agenda – measures of population health

A national research agenda should be developed and funded to determine those measures of the population's health and influences on the population's health for which state and local data are required in order to develop valid and reliable estimates, and to determine those measures from which state and local area estimates can be reliably imputed from national data or data available from other locales.[64] This research agenda should provide for the development of such methods that could be implemented by agencies with varying levels of health statistics expertise.[65]

GP recommendation 6.2: Data set development – geocoding

All ongoing health statistics data sets should be geocoded to at least the census block group level, in order to facilitate geographic aggregation and disaggregation of data.[66]

[62]Centers for Disease Control and Prevention, "Public Health Data Standards Consortium," National Center for Health Statistics, Hyattsville, Maryland, June 2001.

[63]Prior recommendations and related models:
a) NHII recommendation 4 for Federal government, NCVHS,. *Information for Health: A strategy for building the National Health Information Infrastructure*, Ibid. page 41, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].
b) NHII recommendation 2 for state and local government, NCVHS, *Information for Health: A strategy for building the National Health Information Infrastructure*, Ibid. page 42, Available from: http://ncvhs.hhs.gov/nhiilayo.pdf [accessed 24 September 2002].

[64]Prior recommendations and related models: "Vision" discussion group, Albuquerque, 1 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/Albuquerque.pdf [accessed 1 September 2002].

[65]Prior recommendations and related models: Recommendation 2.1.1, source: Australian Institute of Health and Welfare and the National Public Health Information Working Group, "improving the use and delivery of public health information–analysis and presentation of information," *National Public Health Information Development Plan*, Australian Institute of Health and Welfare, Canberra, 1999, page v.

[66]Prior recommendations and related models:
a) Recommendation on geographic infrastructure, source: Canadian Institute for Health Information, *Roadmap Initiative . . . Launching the Process*, Canadian Institute for Health Information, Ottawa, March 2000, pp. 16-17.
b) "Vision" discussion group, New Orleans, 22 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/NewOrleans1.pdf [accessed 1 September 2002].

GP recommendation 6.3: Data set development – topic-specific question modules

Federal agencies should develop and validate topic-specific question modules for population-based surveys, drawing upon the collaborative experiences of the Behavioral Risk Factor Surveillance System, in order to support state and local agencies' development of small-area health statistics.

Guiding Principle 7. Integrated, streamlined data collection for multiple purposes.

GP recommendation 7.1: Data set development – assessing ongoing data sets

Ongoing data collection systems should be regularly assessed to identify overlap and duplication in samples, topics, and data sources. For areas of overlap and duplication, the most appropriate data source and data collection mode should be identified, and redundancy should be eliminated in data collection.[67]

GP recommendation 7.2: Data set development – integrated state collection systems

Building upon mechanisms established under Guiding Principle 1, efforts within individual states to develop integrated data collection systems combining presently discrete systems should be strongly supported and funded. These integrated systems will contain individual-based, longitudinal data from presently discrete reporting, case management, administrative, bill payment, and health statistics data collection systems. Where possible, common portals, business rules, and coordinated databases should be developed across presently discrete data collection systems.[68] Coordination of Federal resources and programs should occur to support development of integrated state data collection systems.

GP recommendation 7.3: Data set development – ongoing multi-purpose data collection systems

State and Federal government agencies should strongly support and fund ongoing multi-purpose data collection systems (such as the National Health Interview Survey, the National Health and Nutrition Examination Survey, the Behavioral Risk Factor Surveillance System, the Vital Statistics Cooperative Program, and state surveys) and

[67]Prior recommendations and related models:
a) Recommendation on a common Web-based reporting architecture, source: California HealthCare Foundation, *Web-based Public Health Reporting in California: A Feasibility Study*, December 2001.
b) Recommendation on National Health Surveillance Infostructure, source: Health Canada, *Building partnerships to improve health surveillance tools and connections across Canada*, 16 September 1999.

[68]Prior recommendations and related models: Rodrian, M., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," San Francisco, 30 October 2000, Available from: http://www.ncvhs.hhs.gov/001030tr.htm [accessed 1 September 2002].

data integration efforts (such as the CDC Assessment Initiative). These efforts, if adequately supported, can meet multiple needs, avoid creation of numerous separate data systems for individual data needs, and serve as "magnets" for further integration and standardization.

Guiding Principle 8. Timely production of valid and reliable health statistics.

GP recommendation 8.1: Data set development – timeliness benchmarks

The national HSPB should develop model standards and benchmarks for timeliness, and individual data systems should adhere to standards for timeliness of release of data and standards for release of regularly scheduled reports. Timeliness standards should mandate release of data and reports no later than within one year of the conclusion of the data year.[69]

GP recommendation 8.2: Research agenda – maximizing timeliness of data release

A research agenda should be developed and funded to develop data collection, collation, analysis, and dissemination methods and technologies that will maximize the timeliness of release of regularly scheduled reports from ongoing data collection systems.

Guiding Principle 9. Appropriate access to and ease of use of health statistics.

GP recommendation 9.1: Data access – Web-based access to data

The massive capacity of the World Wide Web should be exploited to provide more ready access to health statistics through:

- An integrated agenda for the further improvement, expansion, and dissemination of Web-based systems that provide access to national, state, and local data on the population's health and influences on the population's health. These systems should be designed for users of differing levels of statistical sophistication and should be built so they can be searched by health topic and not solely by data set.[70]

[69]Prior recommendations and related models: Office of Management and Budget, "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies," *Federal Register,* 24 September 2001 and 1 October 2001.

[70]Prior recommendations and related models:
a) Friedman, D., et al, "Accessing public health information through interactive systems: lessons learned and future directions," *Public Health Reports*, 2001;116:132-147.
b) Recommendation 2.2.1, improving the use and delivery of public health information–information delivery and access, source: Australian Institute of Health and Welfare and the National Public Health Information Working Group, *National Public Health Information Development Plan*, Australian Institute of Health and Welfare, Canberra, 1999, page v.

- A World Wide Web health statistics locater portal for data on the population's health and influences on the population's health, together with a Web-based meta-data directory containing essential information for each set of data, including descriptions of samples, data collection modes, statistical methods, periodicity, questions or items, and caveats about use.[71]

GP recommendation 9.2: Research agenda – research data centers

An expanded program of Research Data Centers should be developed to provide Center-approved population health researchers pursuing approved research projects with controlled access to potentially identifiable individual-level data on the population's health and influences on the population's health at designated sites, building on models developed by the Census Bureau and NCHS.

GP recommendation 9.3: Data access – best practices

Mechanisms should be developed to disseminate throughout the health statistics enterprise examples of best practices, new methods, and new technologies for maximizing appropriate access to and ease of use of health statistics.[72]

Guiding Principle 10. Continuous evaluation of the scientific integrity, accuracy, and timeliness of health statistics and of the ability of the health statistics enterprise to fulfill its mission.

Recommendations GP 1.1 through 1.6 also address this principle.

GP recommendation 10.1: Evaluation – examples of health statistics uses

The HSPB, and the Planning Boards of each state, should systematically compile feedback on the practical applications of health statistics to policy, research, and public health practice in order to develop a more complete picture of the usefulness of statistics and the ability of the health statistics enterprise to meet needs. This information should be published and disseminated to providers of data, policymakers, and the public. The information can help encourage participation in health statistics systems, make clear the benefits of these systems in debates over resources and privacy, and highlight innovative uses of data that might be replicated elsewhere.

[71]Prior recommendations and related models: Spar, E., letter to NCVHS, 5 October 2001.

[72]Prior recommendations and related models: Upshaw, V., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," Research Triangle Park, 20 November 2000, Available from: http://www.ncvhs.hhs.gov/001120tr.htm [accessed 1 September 2002].

GP recommendation 10.2: Evaluation – NCHS adherence to NAS guidelines

The NCHS Board of Scientific Counselors (Recommendation GP 1.2) should periodically conduct a formal evaluation of the extent to which NCHS adheres to the attributes of a Federal statistical agency, as outlined by the Committee on National Statistics of the National Academy of Sciences.[73]

Recommendations for moving toward the adoption of the overarching conceptual framework

CF recommendation 1: Data set development – measurement of health

The health statistics enterprise, under the leadership of the national HSPB, should examine ongoing data collection systems in order to assess the extent to which functional status and well-being, including mental health, are being measured. Based upon that assessment, measures of functional status and well-being should be incorporated as identified to fill gaps. As needed, research should be conducted on what the most appropriate measures would be and how to operationalize them.[74]

CF recommendation 2: Data set development – data sets for community analyses

A manual and a "tool kit" of data sets for use in community-level analyses should be prepared for use by staff of state and Federal agencies with health statistics responsibilities. The tool kit should emphasize data sets that provide information on contextual variables and community-level attributes.

CF recommendation 3: Data set development – portraits of influences on the population's health throughout the life cycle

Priority should be placed by state and Federal agencies on the development of portraits of the population's health throughout the life cycle. Data sets to provide such portraits should only be constructed after the development of privacy impact statements as called for in GP recommendation 3.1. The constructed data sets should be built and operated with safeguards such as those recommended in the U.S. General Accounting Office's report on *Record Linkage and Privacy: Issues in Creating New Federal Research and*

[73]National Research Council, Martin, M., Straf, M., and Citro, C., ed., *Principles and Practices for a Federal Statistics Agency,* 2nd edition, National Academy Press, Washington, DC, 2001.

[74]Prior recommendations and related models:
a) National Committee on Vital and Health Statistics, *Classifying and Reporting Functional Status*, July 2001.
b) "Vision" discussion group, Albuquerque, 1 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/Albuquerque.pdf [accessed 1 September 2002].
c) National Committee on Vital and Health Statistics, Expert Discussion Group, *Developing A 21st Century Vision for Health Statistics*, Washington, D.C., March and May 1999.

Statistical Information.[75] Constructed data sets may link: administrative health data sets longitudinally on a person basis; data on births, deaths, marriages, and divorces, longitudinally on a person and family basis; multiple surveillance data sets longitudinally on a person basis; survey data sets longitudinally on a person, family, and household basis; and data collected through multiple modes (such as surveys, administrative health data collection systems, surveillance systems, and vital records) combined into single data sets longitudinally on a person basis.[76]

To complement constructed longitudinal data sets, state and Federal agencies should place greater priority on developing and implementing person-based, longitudinal surveys to better understand the various influences on the population's health throughout the life cycle.[77] An example of this type of survey is the congressionally mandated National Children's Study, a multi-agency Federal effort being coordinated by the National Institute for Child Health and Human Development of the National Institutes of Health. This study will examine the effects of environmental influences on the health and development of more than 100,00 children, following them from before birth until age 21.[78]

CF recommendation 4: Data access – regularly released health statistics reports

Regularly released health statistics reports should be enhanced to reflect an overarching conceptual framework of influences on the population's health:

- Models of community and multilevel analyses including both individual-level

[75]United States General Accounting Office, *Record Linkage and Privacy: Issues in Creating New Federal Research and Statistical Information*, Washington, D.C, April 2001, GAO-01-126SP.

[76]Prior recommendations and related models:
a) Canadian Institute for Health Information, *Health Information Roadmap: Responding to Needs*, Canadian Institute for Health Information, Ottawa, 1999, pp. 8-9.
b) Fielding, J., NCVHS public hearing testimony on: "21st Century Vision for Health Statistics," San Francisco, 30 October 2000, Available from: http://www.ncvhs.hhs.gov/001030tr.htm [accessed 1 September 2002].
c) "Vision" discussion group, New Orleans, 22 March 2000, Available from: http://www.ncvhs.hhs.gov/hsvision/NewOrleans1.pdf [accessed 1 September 2002]

[77]Prior recommendations and related models: Recommendation for comprehensive national longitudinal health study, source: National Research Council, Kalsbeek, W. and Perrin, E., "Some new directions," *Toward a Health Statistics System for the 21st Century: Summary of a Workshop*, Perrin, E., Kalsbeek, W., and Scanlan, T., ed., National Academy Press, Washington, D.C., 2001, pp. 64-66.

[78]The National Children's Study, Available from: http://nationalchildrensstudy.gov/ [accessed 1 September 2002].

and community-level data should be developed for regularly released health statistics reports.[79]

- Analyses at multiple geopolitical and population levels (e.g., state, county, community) should be incorporated into regularly released health statistics reports.[80]
- Relationships between the population's health and influences on the population's health should be explored in regularly released health statistics reports.[81]
- Areal linkage of health data to the geographic, demographic, and environmental data that are publicly available from various state and Federal agencies should be included in regularly released health statistics reports; e.g., health data can be linked to the data from the Environmental Protection Agency, Census Bureau, and U.S. Geological Survey that are included in LandViewV.[82]

CF recommendation 5: Training – population health

National organizations should provide intensive teaching/learning sessions focusing on concepts related to the population's health and influences on the population's health at professional meetings attended by staff of state and Federal agencies with health statistics responsibilities, such as meetings of the National Association for Public Health Statistics and Information Systems, the National Association of Health Data Organizations, and the Association of Maternal and Child Health Programs.

[79]Prior recommendations and related models:
a) Diez-Roux, A., "Bringing context back into epidemiology: variables and fallacies in multilevel analysis," *Am J of Public Health.* 1998;88:216-222.
b) Susser, M., "Does risk factor epidemiology put epidemiology at risk? Peering into the future," *J Epidemiology and Community Health*. 1998;58:608-611.
c) Susser, M. and Susser, E., "Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology," *Am J Public Health.* 1996;86:674-677.
d) McKinlay, J. and Marceau, L., "To boldly go . . ." *Am J Public Health.* 2000;25-33.

[80]Prior recommendations and related models: McMichael, A., "Persons, populations, and planets," *Epidemiology*, Volume 6, Number, 6, November 1995, pp. 633-635.

[81]Prior recommendations and related models:
a) National Research Council. *New Horizons in Health: An Integrative Approach*. Burton H. Singer and Carol D. Ryff, Editors. Board on Behavioral, Cognitive, and Sensory Sciences. Washington, D.C.: National Academy Press, 2001.
b) Office of Behavioral and Social Sciences Research, National Institute of Health, "Report of a conference held on June 27-28 2000," *Toward Higher Levels of Analysis: Progress and Promise in Research on Social and Cultural Dimensions of Health*, Available from: http://obssr.od.nih.gov/Conf_Wkshp/higherlevel/ [accessed 7 January 2002].
c) Canadian Institute for Health Information, *Health Information Roadmap: Responding to Needs*, Canadian Institute for Health Information, Ottawa, 1999, pp. 10-11.
d) Health Canada, *The Population Health Template: Key Elements and Actions That Define A Population Health Approach*, July 2001 draft.
e) Frohlich, K. and Potvin, L., "Collective lifestyles as the target for health promotion," *Can J Public Health.* 1999;90:S11-S13.

[82]LandView is a desktop mapping system that incorporates database extracts from the Environmental Protection Agency, the U.S. Census Bureau, and the U.S. Geological Survey. See *Landview*, Available from: http://landview.census.gov [accessed 1 September 2002].

CF recommendation 6: Research agenda–causal relationships

The Federal government should support research into how health statistics can be used more effectively to understand the causal relationships between 1) specific influences on health and functionality and well-being in general and 2) specific high priority diseases (such as lung cancer, breast cancer, childhood asthma, depression) in particular.

Summary of recommendations

Recommendations on data access

- Use the World Wide Web to Improve access to national, state, and local data on the population's health and influences on the population's health. [GP 9.1]
- Disseminate throughout the health statistics enterprise the best methods for maximizing appropriate access to and ease of use of health statistics. [GP 9.3]
- Enhance regularly released health statistics reports to reflect an overarching conceptual framework of influences on the population's health. [CF 4]

Recommendations on data set development

- Develop systems to actively monitor the population's health and potential influences on the population's health in order to identify emerging problems. [GP 4.1]
- Develop mechanisms to rapidly modify the samples, data collection strategies, and data collection instruments of ongoing data collection systems in response to emerging issues. [GP 4.2]
- Geocode all ongoing data sets that feed the health statistics enterprise to the census block group level. [GP 6.2]
- Develop and validate topic-specific question modules for population-based surveys to support state and local agencies' development of small-area health statistics. [GP 6.3]
- Identify duplicative data collection efforts and eliminate redundancy. [GP 7.1]
- Support and fund efforts within individual states to integrate or better coordinate, as appropriate, discrete data systems. [GP 7.2]
- Support and fund ongoing multi-purpose data collection systems and data integration efforts. [GP 7.3]
- Develop and adopt benchmarks for timeliness of release of data and reports derived from those data. [GP 8.1]
- Assure that appropriate measures of functional status and well-being are included in ongoing systems that are a part of the health statistics enterprise. [CF 1]
- Prepare and provide to staff of state and Federal agencies with health statistics responsibilities a manual and a "tool kit" of data sets for use in community-level analyses. [CF 2]
- Develop person-based, longitudinal data sets and surveys in order to develop portraits of influences on the population's health throughout the life cycle. [CF 3]

Recommendations on data standards

- Harmonize the CDC public health conceptual data model with the overarching conceptual framework for the health statistics enterprise in order to facilitate the development of standards. [GP 5.1]
- Adopt or, if necessary, develop standards for data elements commonly used in all methods of data collection, for electronically transmitting data, for presenting and disseminating data, and for providing electronic access to data. [GP 5.2]
- Provide guidance for implementing and auditing the use of standards by the health statistics enterprise. [GP 5.3]
- Contribute to the national standards setting process. [GP 5.4]

Recommendations on enterprise structure

- Assign overall responsibility for health statistics leadership within DHHS to a reconstituted National Center for Health Statistics, supported by a strong Board of Scientific Counselors. [GP 1.1]
- Establish a national *Health Statistics Planning Board* with overall responsibility for planning and agenda setting for health statistics in the United States. [GP 1.2]
- Assign overall responsibility for health statistics planning and coordination activities within each state's government to a single state agency. Establish a *Health Statistics Planning Board* in each state with overall responsibility for planning and agenda setting for health statistics within the state. [GP 1.3]

Recommendations on evaluation

- Compile feedback on the practical applications of health statistics to policy, research, and public health practice and disseminate these findings to providers of data, policymakers, and the public. [GP 10.1]
- Conduct periodic evaluations of NCHS' adherence to the National Academy of Sciences' attributes of a Federal statistical agency. [GP 10.2]

Recommendations on privacy policy

- Develop a code of fair information practices for the health statistics enterprise. [GP 3.1]
- Develop guidance on the application of the HIPAA Standards for Privacy of Individually Identifiable Health Information to health statistics. [GP 3.2]
- Develop model state and Federal privacy, confidentiality, and security statutes for government-sponsored health statistics. [GP 3.3]
- Develop a toolbox of privacy, confidentiality, and security best practices for use throughout the health statistics enterprise. [GP 3.4]

Recommendations on a research agenda

- Develop and fund a research agenda for the study of methods and technologies to enhance the privacy, confidentiality, and security of health statistics. [GP 3.6]
- Develop and fund a research agenda to explore new data collection strategies that can rapidly and flexibly provide data on emerging influences on the population's health; assess the validity and reliability of items used in key ongoing data collection systems; and estimate any loss in accuracy from early publication of provisional, incomplete data from ongoing data collection systems. [GP 4.3]
- Develop methods to validly and reliably estimate important indicators of the health, and of the influences on the health, of state and local populations. [GP 6.1]
- Develop and implement methods and technologies that will maximize the timeliness of release from ongoing data collection systems of regularly scheduled reports. [GP 8.2]
- Provide controlled access to potentially identifiable individual-level data on the population's health and influences on the population's health. Such access would be provided, at designated sites, to researchers pursuing approved projects. [GP 9.2]
- Support research into how health statistics can be used more effectively to understand the causal relationships between specific influences on health and functionality and well-being in general and specific high priority diseases. [CF 6]

Recommendations on training

- Expanded graduate, in-service, and continuous training opportunities should be developed and offered in focusing on all elements of the health statistics enterprise. [GP 1.4]
- Develop a model training agenda and training materials for use throughout the health statistics enterprise focusing on fair information practices, privacy, confidentiality, and security. [GP 3.5]
- Provide training about concepts related to the population's health and influences on the population's health at professional meetings attended by staff of state and Federal agencies with health statistics responsibilities. [CF 5]

Glossary

Aggregate measures are those community attributes that are derivable from the attributes of individual members of the community, such as the infant mortality rate. Aggregate measures can represent average behaviors or attributes as well as the distribution of behaviors or attributes.

Collective lifestyles are "not just the behaviors that people engage in, but rather the relationship between people's social conditions and their behaviours. . . the idea of collective lifestyles proposes that this relationship between social conditions and behaviour is a collective experience, and therefore, may have similar influences on those who partake in this experience."[83]

A **community** is "a unit, generally geographically defined, which is the locus of basic political and social responsibility and in which everyday social interactions involving all or most of the spectrum of life activities of the people within it takes place."[84]

Cultural context is the interrelated "body of customary beliefs, social forms, and material traits constituting a distinct complex of tradition" of a community.[85] The cultural context includes, but is not limited to, norms, values, racism, and sexism.

A **norm** is "an ideal standard binding upon the members of a group and serving to guide, control, or regulate proper and acceptable behavior"[86]

A **value** is "something (as a principle, quality, or entity) intrinsically valuable of desirable."[87]

Sexism is "interpersonal and institutional practices whereby members of dominant gender groups (typically men) accrue privileges by subordinating other gender groups (typically women) and justify these practices via ideologies of innate superiority, difference, or deviance."[88]

Racism is interpersonal and institutional practices whereby members of one racial group accrue privileges by subordinating other racial groups and justify these practices via ideologies of innate superiority, difference, or deviance.[89]

[83]Frohlich, K. and Potvin L., "Collective lifestyles as the target for health promotion," *Can J Public Health.* 1999;90: S13.

[84]Starfield, B., "Basic concepts in population health and health care," *J Epidemiology and Community Health.* 2001;55(7):452.

[85]Webster's Third International Dictionary of the English Language, Unabridged, Springfield, Massachusetts: Merriam-Webster Inc., Publishers, 1993, "culture," definition 5b, p. 552.

[86]Webster's Third International Dictionary of the English Language, op cit., "norm," definition 3, p. 1540.

[87]Webster's Third International Dictionary of the English Language, op cit., "value," definition 8, p. 2530.

[88]Krieger, N., "A glossary for social epidemiology," *J Epidemiology and Community Health*, 2001;55(10):694.

[89]Adapted from Krieger, N., "A glossary for social epidemiology," *J Epidemiology and Community Health.* 2001: 55(10):694.

Ecological measures are those community attributes that are not derivable from the attributes of individual members of the community (e.g., public policies or environmental attributes).

Health services can be characterized through structure, process, access, and use.

Structure refers to those "aspects of the design of health services that influence the way in which services are delivered. These include the numbers and types of personnel and staff; the way in which these personnel are organized to do their work; the nature and extent of the facility and equipment; the range of services offered; systems of management and amenities; accessibility; mode of financing; the way in which the eligible population is determined and enumerated; and the mechanisms of governance and administrative decision making."[90]

Processes of health services care are "aspects of the behavior or performance of the health care system or health facility and the nature of participation on the part of people it serves. These consist of two parts: behaviors of the professionals and participation of the people who relate to the system or facility. Behaviors of professionals include the spectrum of clinical care involving problem (or needs) recognition, the diagnostic process, the recommendation of treatment or management, and appropriate follow up. Participation of people who relate to the system includes utilization, understanding of recommendations, satisfaction with services rendered, and participation in decisions concerning strategies for management."[90]

Utilization is defined as the "experience of people as to their receipt of health care services of different types."[90]

Access is defined as "the perceptions and experiences of people as to their ease in reaching health services or health facilities in terms of location, time, and ease of approach."[90]

Accessibility is defined as "aspects of the structure of health services or health facilities that enhance the ability of people to reach a health care practitioner, in terms of location, time, and ease of approach."[90]

Health statistics are numerical data that characterize the health of a population and the influences and actions that affect the health of a population.

The **health statistics enterprise** consists of the infrastructure and the activities or processes necessary to produce health statistics. This includes the organizations and individuals responsible for collecting data that become health statistics; aggregating and compiling those data into health statistics; analyzing and translating health statistics; and evaluating the health statistics enterprise in order to improve it.[91]

[90]Starfield, B., "Basic concepts in population health and health care," *J Epidemiology and Community Health*. 2001;55(7): 452.

[91]See also the definition and explanation of the health statistics enterprise on page 5.

Political context is determined by public policy, political culture, and political enfranchisement and participation.

Public policy is "a guide to action at any level of government to change what would otherwise occur, a decision about amounts and allocations of resources: the overall amount is a statement of commitment to certain areas of concern; the distribution of the amount shows the priorities of decision making. Policy sets priorities and guides resource allocation."[92]

Political culture is the "integrated pattern of human behavior [relating to politics and political participation] that includes thought, speech, action, and artifacts and depends upon [hu]man's capacity for learning and transmitting knowledge to succeeding generations"[93]

Political enfranchisement and participation refers to is the extent to which population subgroups are allowed to and do engage in politics.

The population's health is "not merely the sum of the health of individuals; it also entails consideration of the nature of the distribution of health throughout the population."[94] Population health has three components: well-being, functional status, and disease.

Well-being is "the extent to which an individual or group is able, on the one hand, to realize aspirations and satisfy needs and, on the other hand, to cope with the interpersonal, social, biological and physical environments."[94]

Functional status is the "ability to engage in the activities of daily living and social life activities."[94]

Disease is medically defined illness, either clinically or subclinically manifested.

[92]Milio, N., "Glossary: healthy public policy," *J Epidemiology and Community Health.* 2001;55(9): 622.

[93]Webster's New Collegiate Dictionary, Springfield, Massachusetts: G & C Merriam Company, publisher, 1981, "culture," definition 5a, p. 274

[94]Starfield, B., "Basic concepts in population health and health care," *J Epidemiology and Community Health*. 2001;55(7): 453.

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95]

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Context				
Natural environment				
Climate and weather (atmosphere)	Weather monitoring stations; satellite data	National Oceanographic and Atmospheric Administration and National Weather Service data files; private companies	National Oceanographic and Atmospheric Administration and National Weather Service data files; private companies	National Oceanographic and Atmospheric Administration and National Weather Service data files; private companies; local television and radio stations
Air quality; water quality; environmental contaminants	Environmental monitoring systems; sentinel surveys and special studies of vectors	Environmental Protection Agency: Aerometric Information Retrieval System and STORET	State environmental protection departments; state Air Quality Control Boards; state health departments	Local water districts; local sewer districts; local sanitation departments; local air quality monitoring programs; local health departments
Geology; topography; vegetative cover; and water resources	Land surveys; satellite images; aerial photographs	U.S. Geological Survey; U.S. Department of Agriculture; U.S. Environmental Protection Agency	State geology departments; state agriculture departments; state environmental protection departments	Local water districts; local land use planning programs
Animal vectors (biosphere)	Field surveys		State health departments; state agriculture departments	Local health departments; local animal control programs

[95]**Sources:**

a) Pearce, N., "Information resources in the United States," *Oxford Textbook of Public Health,* Vol. 2, Detels, R., Holland, W., McEwen, J., and Omenn, G., editors, Oxford University Press, New York, 1997, pp. 435–450;

b) U.S. Department of Health and Human Services. *Tracking healthy people 2010,* DHHS, Washington, D.C., 2000;

c) Health and Human Services, *Part C: Major Data Sources for Healthy People 2010,* available from: http://www.health.gov/healthypeople/Document/html/tracking/THP_PartC.htm [accessed 1 September 2002];

d) Parrish, R., and McDonnell, S., "Sources of health-related information," *Principles and practice of public health surveillance,* Teutsch, S., and Churchill, E., ed.,. Oxford University Press, New York, 2000, pp. 49–54;

e) National Research Council, Institute of Medicine, *Integrating Federal statistics on children: report of a workshop.* NAS Press, 1995.

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Cultural context				
Norms & and values		General Social Survey		
Racism		General Social Survey		
Sexism				
Competition or cooperation				
Political context				
Public policies & and laws (social, economic, health, environment)	Compilations and indexes	Thomas—U.S. Congress on the Internet	National Conference of State Legislatures	
Political culture		General Social Survey		
Differential political enfranchisement or participation				
Community resources				
Biological characteristics				
Population (number, geographic distribution)	Census and periodic population surveys	Decennial census; Current Population Survey	Decennial census; state vital statistics system;	Decennial census
Population (age and sex distribution)	Census and periodic population surveys	Decennial census; Current Population Survey	Decennial census; state vital statistics system;	Decennial census
Biological composition (e.g., genetic endowment, immune status)	Health care records; health examination surveys (e.g., hypertension, immune status);	Human genome project; research studies; outbreak investigations; National Health and Nutrition Examination Survey		
Fertility and natality (rate, geographic and demographic distribution, and temporal trends)	Vital registration; health care records; periodic population or health care provider surveys; periodic surveys of abortion providers	National Vital Statistics System; CDC Abortion Surveillance System; National Survey of Family Growth; Alan Guttmacher Institute	State vital statistics system; National Vital Statistics System; CDC Abortion Surveillance System	State vital statistics systems; National Vital Statistics System

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Built environment				
Land use (deforestation, farming)	Land surveys; satellite images; aerial photographs; population surveys	U.S. Department of Agriculture; Decennial census;	State agriculture departments	
Urbanization (type, location, and extent); and housing		Department of Housing and Urban Development	Decennial census	Local land use planning programs
Workplaces	U.S. population surveys; business and employer surveys; workplace regulatory inspections	Bureau of Labor Statistics; Occupational Safety and Health Administration; Mine Safety and Health Administration; Census Bureau	Current Population Survey; County Business Patterns; Census of Industrial Sectors; Census of Agriculture; state labor departments	
Schools				
Transportation infrastructure	Land surveys; satellite images; aerial photographs	U.S. Department of Transportation	State transportation departments	Local road and highway departments
Communication infrastructure (access to health information)	Population surveys; government agencies regulatory activities; and mandated reporting by private organizations	Federal Communications Commission; Census Bureau; private telecommunication companies	Private telecommunication companies	Private telecommunication companies
Social				
Social support				
Social influence				
Social cohesion				
Education (literacy)				

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Economic				
Economic development and equity (e.g., income distribution, employment, access to material resources)	Population surveys; national financial records; World Bank;	Department of Commerce; Department of Labor, Bureau of Labor Statistics;	State economic development departments of economic development	Local Chamber of Commerce
Work environment	Population surveys; business and employer surveys; workplace regulatory inspections	Bureau of Labor Statistics; Occupational Safety and Health Administration; Mine Safety and Health Administration; Census Bureau; National Council on Compensation; National Agricultural Workers Survey	Current Population Survey; County Business Patterns; Census of Industrial Sectors; Census of Agriculture; state labor departments	
Population health programs				
Air quality; water supply and quality; waste disposal	Environmental laws and regulations; data collection specific programs during environmental doing cleanup; mandated reporting of locations of major sources of pollution; research studies of pollutants and their effects; environmental monitoring systems; sentinel surveys and special studies of vectors	Environmental Protection Agency: Aerometric Information Retrieval System and STORET	State environmental protection departments; state air quality control boards; state health departments	Local water district; local sewer district; local sanitation department; local air quality monitoring programs; local health departments

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Personal health practices				
Prevalence; geographic and demographic distribution	Populations surveys (e.g., behaviors, activities); observations of behavior; health examination surveys (e.g., body fat, nutritional status); qualitative research; behavioral risk factor surveys; exit surveys; focus groups	National Health Interview Survey; Youth Risk Behavior Survey; National Household Survey on Drug Abuse; Monitoring the Future Study; 1999 National Worksite Health Promotion Survey; Mothers' Survey, Abbott Laboratories, Inc., Ross Products Division	State departments of health; state departments of education; state Behavioral Risk Factor Surveys; Youth Risk Behavior Survey; School Health Policies and Programs Study; State Tobacco Activities Tracking and Evaluation System	Youth Risk Behavior Survey

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Health services				
Providing effective health care (Number and types of persons individuals and facilities; geographic distribution; type of care given; number of people that can be served; utilization of health care)	Government health reports; licensure of health professionals; surveys of health facilities and health care providers; populations surveys; studies of health care quality and access to care; accreditation and licensing of health care providers and facilities; reporting systems from professional associations; census of local health departments; reporting systems from for physicians who treat end-stage renal disease	National Health Care Survey; American Hospital Association Annual Survey of Hospitals; American Medical Association Physician Masterfile; Healthcare Cost and Utilization Project; National Medical Expenditure Survey; Joint Commission on the Accreditation of Health Care Organizations; Clinical Preventive Services Guidelines; National Immunization Survey; National Profile of Local Health Departments; U.S. Renal Data System; American Association of Colleges of Osteopathic Medicine reports on osteopathic medical education; American Association of Colleges of Pharmacy enrollment report; American Dental Association surveys of dental educational institutions; Annual Census of Hospitals; Association of American Medical Colleges reports on enrollment; Association of Schools and Colleges of Optometry data on enrollment; Association of Schools of Public Health data on schools of public health; National Health Maintenance Organization Census; National League of Nursing annual survey of schools of nursing	National Immunization Survey	National Profile of Local Health Departments; U.S. Renal Data System

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Population health				
Disease				
Mortality (number, geographic and demographic distribution, temporal trends, causes); and life expectancy	Vital registration; health care records; periodic population or health care provider surveys; "verbal autopsy"	National Vital Statistics System; Bureau of Labor Statistics; National Institute of Occupational Safety and Health	State vital statistics systems; state medical examiners and coroners; Census of Fatal Occupational Injuries	Local (county) medical examiners and coroners
Morbidity (number, nature and severity, geographic and demographic distribution, temporal trends, causes)	General: registers of visits to health care facilities[96]; electronic health care records; administrative or financial data derived from health care records (e.g., insurance claims, data collected by payers from health care providers in order to process payment for providing health care, and reports required by government or regulatory agencies); population surveys; Hospitalizations: hospital records; hospital discharge surveys;	General: databases maintained by managed care organizations; MEDSTAT research databases and other commercial health-care data sets; national Medicaid data files; Medicare data files Hospitalizations: National Hospital Discharge Survey; national Medicaid data files; Healthcare Cost and Utilization Project; Medical Expenditure Panel Survey Ambulatory: National Ambulatory Medical Care Survey and National Hospital Ambulatory Medical Care Survey (NCHS); HIV/AIDS Surveillance System;	General: databases maintained by managed care organizations; MEDSTAT research databases and other commercial health-care data sets; state Medicaid data files; Medicare data files Hospitalizations: state hospital discharge surveys; state Medicaid data files; Healthcare Cost and Utilization Project; Medical Expenditure Panel Survey	General: databases maintained by Managed Care Organizations; MEDSTAT research databases and other commercial health-care data sets; state Medicaid data files; Medicare data files Hospitalizations: state hospital discharge surveys; state Medicaid data files; Healthcare Cost and Utilization Project; Medical Expenditure Panel Survey

[96]Type of health care facility visited is often used as a surrogate for severity.

Table 1. Influences on the health of populations, typical approaches for gathering information on these influences, and systems that gather this information in the United States at the national, state, and local levels[95] (continued)

Influences on health	Typical approaches	Examples of systems: national level	Examples of systems: state level	Examples of systems: local level
Morbidity (continued)	Specific diseases: disease notifications; reporting of sentinel disease by health-care providers; required reporting of epidemic diseases ("disease notifications"); registries of patients with specific diseases kept by health care providers or facilities, or health agencies; surveys of employers (occupational injury and illness records; laboratory reporting	Specific diseases: National Notifiable Disease Surveillance System; Drug Abuse Warning Network; National Electronic Injury Surveillance; Annual Survey of Occupational Injuries and Illnesses;National Agricultural Workers Survey	Ambulatory: Specific diseases: National Notifiable Disease Surveillance System; state notifiable disease programs; state cancer registries; state birth defects registries; Drug Abuse Warning Network; National Electronic Injury Surveillance	Ambulatory: Specific diseases: state notifiable disease programs; state cancer registries; state birth defects registries; Drug Abuse Warning Network; National Electronic Injury Surveillance
Function				
	Population surveys (self-reported health status)	National Health Interview Survey; National Health and Nutrition Examination Survey; National Crime Victimization Survey	State Behavioral Risk Factor Surveys; State and Local Area Integrated Telephone Survey	State and Local Area Integrated Telephone Survey
Well-being				

Table 2. Programs and activities that contribute to the Health Statistics Cycle in the United States at the national, state, and local levels.

Step in Cycle	Examples of programs and activities: national level	Examples of programs and activities: state level	Examples of programs and activities: local level
Define data needs	Legislation and Congressional support organizations (e.g., Congressional Research Service, General Accounting Office, Congressional Budget Office); administration policy priorities; Federal agency policy and regulatory agendas; Operational and research needs of program managers of Federal program operations; advisory bodies (e.g., National Committee on Vital and Health Statistics); findings of review and study panels (e.g., National Academy of Sciences and Institute of Medicine); reports of commissions; scientific literature	State legislatures; executive agency policies; managers of state program operations; advisory bodies operational and research needs of program managers	County and city councils; local executive agency policies; managers of local program operations; Turning Point grantees
Specify necessary data attributes	Centers and programs within CDC, including the National Center for Health Statistics; Agency for Healthcare Quality and Research; other Federal statistical agencies; other HHS components	State Centers for Health Statistics; state epidemiologists; state chronic disease epidemiologists; Robert Wood Johnson Foundation-funded State Health Access Data Assistance Center; Urban Institute-Assessing New Federalism	
Identify appropriate data sources	FedStats, a gateway to statistics from Federal agencies[97]	State Centers for Health Statistics; state epidemiologists; state chronic disease epidemiologists	
Collect, aggregate, and compile data	Multiple centers and programs within CDC, including the National Center for Health Statistics; Agency for Healthcare Quality and Research; other Federal statistical agencies; other HHS components; foundation-supported entities organizations such as the Center for Studying Health Systems Change	State Centers for Health Statistics; state epidemiologists; state chronic disease epidemiologists	

[97] www.fedstats.gov/

Table 2. Programs and activities that contribute to the Health Statistics Cycle in the United States at the national, state, and local levels. (continued)

Step in Cycle	Examples of programs and activities: national level	Examples of programs and activities: state level	Examples of programs and activities: local level
Analyze and interpret data	Centers and programs within CDC, including the National Center for Health Statistics; Agency for Healthcare Quality and Research; other Federal statistical agencies; other HHS components; foundation-supported organizations such as the Center for Studying Health Systems Change	State Centers for Health Statistics; state epidemiologists; state chronic disease epidemiologists	
Communicate and translate findings [statistics]	FedStats, a gateway to statistics from Federal agencies[98]	State Centers for Health Statistics; state epidemiologists; state chronic disease epidemiologists; state health department Web sites (e.g., Missouri's MICA, Massachusetts' MassCHIP)	County and other local health department Web sites
Evaluate extent to which needs are met	DHHS Data Council; Committee on National Statistics, National Academy of Sciences; NCVHS		
Integrating statistics	DHHS Data Council; NCVHS; standards development organizations; CDC's NEDSS program; Public Health Standards Data Consortium		

[98] www.fedstats.gov/

Appendix: National Committee on Vital and Health Statistics, 2002 Membership

(note: Members of the Workgroup on Health Statistics for the 21st Century are noted in **bold text**.

John R. Lumpkin, M.D., M.P.H., Chair
Director
Illinois Dept. of Public Health
Springfield, Illinois

Brady A. Augustine, M.S.
Corporate Director for Special Projects
Senior Biostatistician and Medical Economist
Gambro Healthcare USA, Inc.
Fort Lauderdale, Florida

Jeffrey S. Blair, M.B.A.
Vice President
Medical Records Institute
Albuquerque, New Mexico

Simon P. Cohn, M.D., M.P.H.
National Director for Health Information Policy
Kaiser Permanente Medical Care Program
Oakland, California

Kathryn L. Coltin, M.P.H.
Director, External Quality and
Data Initiatives
Harvard Pilgrim Health Care
Wellesley, Massachusets

John W. Danaher, M.D.
Litchfield, Connecticut

Daniel Friedman, Ph.D. (Workgroup Chair)
Assistant Commissioner
Bureau of Health Statistics,
Research and Evaluation
Massachusetts Department of Public Health
Boston, Massachusetts

Richard K. Harding, M.D.
Professor of Clinical Psychiatry
and Pediatrics
Vice Chairman, Clinical Services
Neuropsychiatry & Behavioral Science
University of South Carolina
Columbia, South Carolina

Eugene J. Lengerich, V.M.D.
Penn State University
Hershey, Pennsylvania

Vickie M. Mays, Ph.D., M.P.H.
Department of Psychology
University of California, Los Angeles
Los Angeles, California

Clement Joseph McDonald, M.D.
Distinguished Professor of Medicine
Indiana University School of Medicine
Director
Regenstrief Institute
Indianapolis, Indiana

Paul Newacheck, Dr. P.H.
Professor of Health Policy and Pediatrics
Institute of Health Policy Studies
School of Medicine
University of California, San Francisco
San Francisco, California

Mark A. Rothstein, J.D.
Herbert F. Boehl Chair of Law and Medicine
Director, Institute for Bioethics, Health Policy
and Law
University of Louisville School of Medicine
Louisville, Kentucky

Edward H. Shortliffe, Ph.D., M.D.
Department of Medical Informatics
Columbia University
New York, New York

Barbara Starfield, M.D., M.P.H.
Distinguished University Professor
and Professor of Health Policy and Pediatrics
School of Hygiene/Public Health
The Johns Hopkins University
Baltimore, Maryland

Kepa Zubeldia, M.D.
President
Claredi
Kaysville, Utah