



Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network

Edward B. Perrin, Jane S. Durch, and Susan M. Skillman, Editors; Panel on Performance Measures and Data for Public Health Performance Partnership Grants, National Research Council

ISBN: 0-309-51804-0, 192 pages, 6 x 9, (1999)

This PDF is available from the National Academies Press at:
<http://www.nap.edu/catalog/6487.html>

Visit the [National Academies Press](http://www.nap.edu) online, the authoritative source for all books from the [National Academy of Sciences](http://www.nap.edu), the [National Academy of Engineering](http://www.nap.edu), the [Institute of Medicine](http://www.nap.edu), and the [National Research Council](http://www.nap.edu):

- Download hundreds of free books in PDF
- Read thousands of books online for free
- Explore our innovative research tools – try the “[Research Dashboard](#)” now!
- [Sign up](#) to be notified when new books are published
- Purchase printed books and selected PDF files

Thank you for downloading this PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, [visit us online](#), or send an email to feedback@nap.edu.

This book plus thousands more are available at <http://www.nap.edu>.

Copyright © National Academy of Sciences. All rights reserved.

Unless otherwise indicated, all materials in this PDF File are copyrighted by the National Academy of Sciences. Distribution, posting, or copying is strictly prohibited without written permission of the National Academies Press. [Request reprint permission for this book](#).

Health Performance Measurement in the Public Sector

Principles and Policies for Implementing an Information Network

Edward B. Perrin, Jane S. Durch, and Susan M. Skillman, *Editors*

Panel on Performance Measures and Data
for Public Health Performance Partnership Grants

Committee on National Statistics

Commission on Behavioral and Social Sciences and Education

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1999

NATIONAL ACADEMY PRESS 2101 Constitution Avenue, N.W. Washington, D.C. 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract No. 282-95-0034 between the National Academy of Sciences and the U.S. Department of Health and Human Services. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. SBR-9709489). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

Library of Congress Cataloging-in-Publication Data

Health performance measurement in the public sector : principles and policies for implementing an information network / Edward B. Perrin, Jane S. Durch, and Susan M. Skillman, editors ; Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, National Research Council.

p. cm.
Includes bibliographical references (p.).
ISBN 0-309-06436-8 (pbk.)

1. Public health—United States—Information services. 2. Public health—United States—Evaluation—Data processing. I. Perrin, Edward. II. Durch, Jane. III. Skillman, Susan M. IV. National Research Council (U.S.). Panel on Performance Measures and Data for Public Health Performance Partnership Grants.

RA423.2 .H43 1999
362.1'0973—dc21 99-6159

Additional copies of this report are available from National Academy Press, 2101 Constitution Avenue, N.W., Box 285, Washington, D.C. 20055. Call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area). This report is also available on line at <http://www.nap.edu>

Printed in the United States of America

Copyright 1999 by the National Academy of Sciences. All rights reserved.

**PANEL ON PERFORMANCE MEASURES AND DATA FOR
PUBLIC HEALTH PERFORMANCE PARTNERSHIP GRANTS**

EDWARD B. PERRIN (*Chair*), Department of Health Services, School of
Public Health and Community Medicine, University of Washington

JOHN C. BAILAR III, Department of Health Studies, University of Chicago

THOMAS J. BARRETT, Mental Health Services, Colorado Department of
Human Services

DAVID W. FLEMING, Center for Disease Prevention and Epidemiology,
Oregon Health Division

V. JOSEPH HOTZ, Departments of Economics and Policy Studies, University
of California, Los Angeles

IRA R. KAUFMAN, University of Medicine and Dentistry of New Jersey,
Department of Environmental and Community Medicine, Robert Wood
Johnson Medical School

JOHN R. LUMPKIN, Illinois Department of Public Health

WILLIAM A. MORRILL, Mathtech, Inc., Princeton, New Jersey

R. HEATHER PALMER, Center for Quality of Care Research and Education,
Harvard School of Public Health

J. SANFORD SCHWARTZ, School of Medicine and the Wharton School,
University of Pennsylvania

MARY E. STUART, Maryland Collaboratory for Business and Health, Depart-
ment of Sociology and Anthropology, University of Maryland, Baltimore
County

CYNTHIA P. TURNURE, Minnesota Department of Health

PAUL J. WIESNER, DeKalb County Board of Health, Decatur, Georgia

GAIL R. WILENSKY, Project Hope, Bethesda, Maryland

JANE S. DURCH, *Study Director*

JEFFREY J. KOSHEL, *Study Director* (through December 1997)

SUSAN M. SKILLMAN, *Senior Research Associate*

TELISSIA M. THOMPSON, *Senior Project Assistant*

COMMITTEE ON NATIONAL STATISTICS
1998-1999

JOHN E. ROLPH (*Chair*), Department of Information and Operations Management, University of Southern California

JOSEPH G. ALTONJI, Institute for Research on Poverty and Department of Economics, Northwestern University

JULIE DaVANZO, RAND, Santa Monica, California

WILLIAM F. EDDY, Department of Statistics, Carnegie Mellon University

WILLIAM KALSBECK, Department of Biostatistics, University of North Carolina

RODERICK J.A. LITTLE, School of Public Health, University of Michigan

THOMAS A. LOUIS, School of Public Health, University of Minnesota

CHARLES F. MANSKI, Department of Economics, Northwestern University

WILLIAM NORDHAUS, Department of Economics, Yale University

JANET L. NORWOOD, Urban Institute, Washington, D.C.

EDWARD B. PERRIN, Department of Health Services, School of Public Health and Community Medicine, University of Washington

PAUL R. ROSENBAUM, Department of Statistics, Wharton School, University of Pennsylvania

FRANCISCO J. SAMANIEGO, Division of Statistics, University of California, Davis

RICHARD L. SCHMALENSEE, Sloan School of Management, Massachusetts Institute of Technology

MIRON L. STRAF, *Director*

ANDREW A. WHITE, *Deputy Director*

Contents

PREFACE	ix
EXECUTIVE SUMMARY	1
1 INTRODUCTION	15
Origins of the Study, 16	
<i>Performance Partnership Grants, 16 / Other Influences, 17</i>	
Charge to the Panel, 18	
Essential Definitions, 19	
<i>Performance Measurement and Related Concepts,</i>	
<i>19 / Categories of Performance Measures, 20</i>	
Phase I: Focus on Selection of Performance Measures, 22	
<i>Use of Measures of Process and Capacity as Well as</i>	
<i>Outcomes, 22 / Guidelines for Selecting Performance</i>	
<i>Measures, 24 / Limitations of a Program-Specific Approach to</i>	
<i>Performance Measurement, 25 / Need to Strengthen State and</i>	
<i>Local Capacity for Data Collection and Analysis,</i>	
<i>26 / Inadvisability of Using Performance Measures Alone for</i>	
<i>Resource Allocation Purposes, 26</i>	
Phase II: Data and Information System Development to Support	
Performance Measurement, 27	
<i>A Vision for a National Health Information Network,</i>	
<i>27 / Critical Issues, 28</i>	
Structure of the Report, 29	
References, 29	

2	PUTTING PERFORMANCE MEASUREMENT IN CONTEXT	30
	Use of Performance Measurement in Accountability Systems, 30 <i>Performance-Based Accountability Systems, 31 / Operation of Performance-Based Systems, 33 / Applications of Performance Measurement, 37</i>	
	Examples of Performance Monitoring and Accountability Systems, 40 <i>Government Performance and Results Act, 41 / Use of Performance Measures in the Maternal and Child Health Services Block Grant Program, 42 / Reporting Requirements for the Temporary Assistance for Needy Families Program, 44 / State Developments in Performance-Based Budgeting, 45 / Health Care Performance Measurement in the Private Sector, 48</i>	
	Conclusions, 52	
	References, 55	
3	PERFORMANCE MEASUREMENT CONSIDERATIONS FOR PUBLICLY FUNDED HEALTH PROGRAMS	59
	Broad Array of Health-Related Services and Service Relationships, 60	
	Measurement Considerations for Population-Based Health Services, 61 <i>Population Health Services, 61 / Monitoring Population Health Services, 62 Monitoring the Infrastructure for Publicly Funded Health Programs, 64</i>	
	Some Performance Measurement Considerations Related to Program-Specific Matters, 66 <i>Environmental Health Programs, 66 / Mental Health Programs, 68 / Substance Abuse Programs, 71</i>	
	Process Guidelines as a Basis for Performance Measurement, 74 <i>Guidelines for Personal Health Services, 75 / Guidelines for Population-Based Health Services, 76 / Research Needs for Practice Guidelines and Performance Measurement, 76</i>	
	Conclusions, 78	
	References, 79	
4	DATA AND INFORMATION SYSTEMS: ISSUES FOR PERFORMANCE MEASUREMENT	83
	Health Data Resources, 84 <i>Registries, 85 / Surveys, 86 / Patient Records and Related Clinical Encounter Data, 87 / Administrative Data, 89 / Claims Data, 90 / Linkage of Data Sets, 91 / Steps Toward Integration of Data Sets, 93</i>	

Assuring the Quality of Data and Data Analysis, 95	
<i>Random Variation and Bias, 96 / Data Management,</i>	
<i>97 / Challenges in Data Analysis, 98</i>	
Developing and Implementing Standards for Data and Data Systems, 103	
<i>Standardization Activities, 103 / Balancing Standardization</i>	
<i>and Change, 108</i>	
Enhancing Performance Measurement through Advances in Information Technology, 109	
<i>Data Collection and Transmission Technologies, 109 / Data</i>	
<i>Management and Analysis, 110 / Computer-Based Patient</i>	
<i>Records, 111 / The Internet and the World Wide Web,</i>	
<i>111 / Limits of Technology, 114</i>	
Privacy, Confidentiality, and Security of Health Data, 114	
Investing in Health Data and Data Systems, 116	
<i>Data and Information Systems, 117 / Technology, 119 /</i>	
<i>Training and Technical Assistance, 120</i>	
Taking a Collaborative Approach to the Development of Health Data and Information Systems, 122	
<i>Meeting the Needs of Many Data Users, 123 / Collaboration in</i>	
<i>the Design and Implementation of Data Systems, 125</i>	
Conclusions, 126	
References, 127	
 5 STRATEGIES FOR SUPPORTING PERFORMANCE MEASUREMENT THROUGH A NATIONAL INFORMATION NETWORK	 132
Conclusions Regarding Performance Partnership Agreements, 133	
A National Information Network for Health-Related Data, 134	
Essential Features of a Health Information Network to Serve Different Levels of Decision Making, 135	
Major Policy Actions Needed, 138	
<i>National Collaboration, 138 / Integration of Data Systems,</i>	
<i>143 / Technology Policy, 144</i>	
Operational Principles, 145	
<i>Involve a Broad Range of Public Agencies, 145 / Build on</i>	
<i>Existing Information Systems, 146 / Implementation Standards,</i>	
<i>147</i>	
Required Investment for an Information Network, 152	
<i>Data Systems, 152 / Training and Technical Assistance, 154</i>	
A National Research Agenda, 156	
Final Observations, 157	
References, 158	

APPENDICES

A	POTENTIAL HEALTH OUTCOME AND RISK STATUS MEASURES	163
B	WORKSHOP AGENDA	168
C	BIOGRAPHICAL SKETCHES	170

Preface

This is the second of two reports from the Panel on Performance Measures and Data for Public Health Performance Partnership Grants. In summer 1995, as part of its response to the need to ensure that public funding of health programs is related to documented program performance, the Office of the Assistant Secretary for Health of the U.S. Department of Health and Human Services (DHHS) requested that the National Research Council (NRC) convene an expert panel to examine and report on the technical issues involved in establishing performance measures in ten substantive program areas. Such measures were to be required as part of the proposed Performance Partnership Grants (PPG) Program, under which each state was to negotiate with DHHS an action plan with performance objectives that would be specific in terms of outcomes, processes, and capacity to be achieved within a 3- to 5-year period.

The panel divided its work into two phases. The objective of the first phase was to develop a report to the Secretary of DHHS on performance measures in specified areas that would be useful to the PPG Program in its first 3-5 years. The report *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health* (National Academy Press, Washington, D.C., 1997) presents the findings and recommendations resulting from the first phase of the panel's work. The focus of the second phase was on recommended improvements to existing local, state, and federal data systems that would facilitate the collection of data for use with existing and proposed performance measures. This two-phase approach was adopted by the panel because of our conclusion that developmental work is needed in both the public and private sectors to adapt, refine, or add to existing data systems to make them more useful in performance measure-

ment, and to delineate more clearly the complex causal links between program processes and outcomes.

Although formal legislation mandating PPGs has not been adopted, various DHHS programs and a growing number of states and communities are moving to monitor and analyze health outcomes on their own. Moreover, there appears to be a growing consensus within the public health, substance abuse, and mental health communities about the value of performance measurement. Indeed, many people believe that the case for increasing, or even maintaining, public funding will depend on documented program performance. The development of performance measures is a continuation of earlier efforts to assess progress toward important public health goals. The broad acceptance of immunization rates and other such measures developed for Healthy People 2000, for example, has been instrumental in the creation of databases and the mobilization of resources in many jurisdictions to assess progress toward those objectives.

The work of this panel was furthered significantly by four regional meetings of state officials and consumers convened by DHHS early in the process—in Portland, Oregon; San Francisco, California; Chicago, Illinois; and Philadelphia, Pennsylvania—and by input from several national associations of state agency administrators (the Association of State and Territorial Health Officials, the National Association of County and City Health Officials, the National Association of State Alcohol and Drug Abuse Directors, and the National Association of State Mental Health Program Directors). Our work during the second phase of the project was aided by the contributions of participants in a workshop held by the panel in July 1997. The workshop agenda and participants are listed in Appendix B. We also wish to thank Winnie Mitchell of the Substance Abuse and Mental Health Services Administration and David Moriarty of the Centers for Disease Control and Prevention, who provided special assistance by arranging several meetings with agency staff members.

The panel was fortunate to have for its deliberations two specially commissioned working papers. “Improving Federal-State Data Collection to Monitor Program Performance” was prepared by the Science and Epidemiology Committee of the Association of State and Territorial Chronic Disease Program Directors and the Council of State and Territorial Epidemiologists. Christopher Maylahn, New York State Department of Health, chaired the group that prepared this paper. “Quality Education Data: Unprecedented Opportunity for a Decade to Build” was prepared by Emerson Elliott, National Council for Accreditation of Teacher Education, and John Ralph, National Center for Education Statistics. Copies of these papers are available from the Committee on National Statistics (2101 Constitution Avenue, N.W., Washington, DC 20418).

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC’s Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institu-

tion in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We thank the following individuals for their participation in the review of this report: Thomas F. Babor, School of Medicine, University of Connecticut Health Center; Maureen M. Henderson, School of Medicine, University of Washington (emeritus); Ed Lichtenstein, Oregon Research Institute, Eugene; Thomas Louis, School of Public Health, University of Minnesota; J. Michael McGinnis, scholar in residence, National Research Council; Barbara J. McNeil, Department of Health Care Policy, Harvard Medical School; Donald M. Steinwachs, School of Hygiene and Public Health, Johns Hopkins University; Bernard Turnock, Community Health Sciences, University of Illinois, Chicago; and Albert W. Wu, Health Services Research Center, Johns Hopkins University.

While the individuals listed above have provided constructive comments and suggestions, it must be emphasized that responsibility for the final content of this report rests entirely with the authoring committee and the institution.

The panel appreciates the assistance of the staff of the Committee on National Statistics (CNSTAT) in preparing this report. We wish to thank in particular Miron Straf, staff director of CNSTAT, for developing the original project design; Telissia Thompson for organizing panel meetings and carrying out innumerable administrative tasks; and Kathleen Saslaw for providing financial management for the study. Several former CNSTAT staff members provided valuable assistance at various stages in the production of the panel's two reports: Anu Das, Theresa Raphael, Michelle Ruddick, and Ashley Bowers. We appreciate the editorial work of Rona Briere and the guidance of Eugenia Grohman, associate director for reports of the Commission on Behavioral and Social Sciences and Education. Thanks also go to Sue Skillman, University of Washington, for her valued assistance to the committee chair. Finally, the panel extends its special thanks to Jeff Koshel, the panel's study director for the first part of the study, and to Jane Durch, for her very capable management of and contribution to the process leading to the production of this second report.

Edward B. Perrin, *Chair*
Panel on Performance Measures and Data for
Public Health Performance Partnership Grants

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

The title is centered and consists of five lines of bold, black, sans-serif text. The first three lines, 'Health', 'Performance', and 'Measurement', are each preceded by a thick, solid gray horizontal bar. The fourth line, 'in the', is not preceded by a bar. The fifth line, 'Public Sector', is also not preceded by a bar. The bars for 'Performance' and 'Measurement' are of equal length, while the bar for 'Health' is slightly shorter.

Health Performance Measurement in the Public Sector

Executive Summary

Publicly funded health programs are being asked to account for their performance, and methods of performance measurement have emerged as essential tools for operationalizing this quest for accountability. A system of performance measurement promises improved documentation of the contributions of public agencies, and can serve as a quality improvement tool by drawing attention to practices shown to promote the achievement of desired outcomes and by identifying areas needing improvement. In fact, informed observers in public health, health policy, health economics, and related subjects believe that we cannot expect public funding to increase or even be maintained at current levels without better documentation of the return on program investments.

In 1995, the U.S. Department of Health and Human Services (DHHS) proposed legislative changes to establish Performance Partnership Grants (PPGs) for a set of federal block grant programs that provide funding to states for public health, substance abuse, and mental health activities.¹ The proposal called for DHHS and each state to negotiate an agreement on program objectives and a set of performance measures to be used to monitor progress toward those objectives over a 3- to 5-year period. PPGs were proposed because it was felt that they would provide a mechanism for increasing state flexibility in the use of federal grant funds while enhancing accountability for progress toward program goals.

¹The specific program areas covered by the original PPG proposal are chronic diseases; sexually transmitted diseases, human immunodeficiency virus infection, and tuberculosis; immunization; mental health; substance abuse; and three areas of special interest to DHHS—sexual assault, disabilities, and emergency medical services.

Although formal legislative requirements for PPGs have not been established, DHHS and state and local health agencies remain very interested in developing agreements through which these performance measurement and performance partnership concepts can be applied.

The Panel on Performance Measures and Data for Public Health Performance Partnership Grants was assembled in fall 1995 at the request of DHHS to assess the state of the art in performance measurement for the specific block grant programs included in the original PPG proposal, recommend measures that could be used to monitor the proposed PPG agreements to be negotiated between each state and the federal government, and recommend steps to improve performance measures and performance measurement for health-related programs.

The panel's work has resulted in two reports, of which this is the second. In its first report, the panel presented an analytic framework for use by states and DHHS in assessing the appropriateness of specific outcome, process, and capacity measures for individual performance agreements. That report also provided examples of health outcome and risk status measures, and related process and capacity measures, that might be used in conjunction with PPGs in the program areas covered by the original proposal.

In this second report, the panel addresses broader data and information system issues that require attention at the federal, state, and local levels to advance the practice of performance measurement for publicly funded health programs. This report broadens the discussion begun in the first report by looking beyond the federal-state PPG framework and beyond the specific program areas covered by the PPG proposal. Although the discussion focuses on the public-sector perspective, closely related private-sector interests and developments in clinical health care are also considered.

PRINCIPLES OF PERFORMANCE MEASUREMENT

Performance measurement is not new, but its increasingly widespread use in both the public sector (e.g., in federal agencies under the Government Performance and Results Act and in state and local government accountability systems) and the private sector (e.g., in accreditation programs for health care organizations) reflects changing attitudes and expectations regarding accountability and management. An attractive feature of performance measurement is that it presents the opportunity to focus attention on defining and using evidence-based best practices to achieve desired outcomes. To define those best practices and select appropriate performance measures, evidence is needed not only on the causal links between interventions and outcomes, but also on how interventions work. This evidence should guide the organization, operation, and improvement of the services, as well as the selection and use of meaningful process and capacity performance measures. When evidence is not yet conclusive, the selection of performance measures should be guided by professional consensus regarding the

relationship between capacity and process factors and health outcomes. Explicit accountability for those processes and intermediate outcomes that are under programmatic control will encourage the further development of standards of practice in publicly funded health programs.

Performance measurement, as it has been implemented in various settings, involves (1) developing an explicit set of goals and objectives and articulating a strategy for achieving them, (2) developing and implementing strategies for measuring performance, and (3) using performance information to improve management practices or resource allocation. However, the panel cautions against using performance measures as the sole basis for causal inferences regarding program performance because of the diversity of factors beyond program activities that affect most health outcomes.

Successful implementation of performance measurement will require substantial and continuing efforts to overcome challenges involved in the development and use of performance measures and data systems. There must be a firm commitment to ongoing research to develop new and better measures, relate these measures to program actions, and evaluate and improve the performance measurement system. Performance measurement should also provide a framework for initiating systematic efforts, especially at the state and local levels, to identify and collect data for appropriate measures of program capacity, processes, and intermediate outcomes. Performance measurement activities will benefit from collaborations among agency staff and others who can bring to bear the perspectives and expertise of a variety of disciplines (e.g., public health, clinical services, statistics, epidemiology, data processing, data management).

The panel concluded that the following principles should guide current efforts to implement performance measurement for publicly funded health programs:

- Performance measurement activities must proceed from clearly defined program goals. The performance measurement process and the information it produces should be viewed as a tool for monitoring and promoting progress toward those goals, not as an end in itself.
- Because health needs and priorities vary across the country, a performance measurement system should promote the development of identifiable sets of measures from which states and communities can select subsets appropriate for the program priorities and strategies they have adopted. Ideally, all of the measures used should be recognized as valid, reliable, and responsive to change and have agreed-upon definitions.
- Performance measurement activities must recognize and meet information needs that may differ in terms of content and quantity for different functions (e.g., program operation, management, policy making, funding). Ideally, measures for specific functional purposes should be linked, conceptually or in practice, to provide a consistent assessment of performance across these different functions.

- Performance measurement must consider the feasibility of data collection and analysis for proposed measures. If appropriate data cannot be produced, the measures will not be useful. Considerations should include the quality and comparability of the available data and the cost of producing those data.
- The performance monitoring system, including the performance measures and the data supporting them, should be evaluated periodically. Such evaluation will help ensure that the system's goals are being met and decrease the likelihood of manipulation or inadvertent promotion of undesired effects, such as reducing services to groups that may be likely to have poor outcomes.
- Performance measurement should be viewed as a developmental activity that will continue to evolve. Measures should be refined or replaced as understanding of the linkages between health outcomes and program activities ("processes") improves, as better sources of data are developed, and as program priorities change. The panel cautions that because performance measurement is a new and largely unfamiliar policy mechanism, it should be tested in the context of goal setting, progress monitoring, and signaling before being used for resource allocation or regulatory purposes. Research and evaluation studies should assess performance measurement's effectiveness as a tool for improving health outcomes and program management.

PERFORMANCE PARTNERSHIP AGREEMENTS

Examination of performance measurement in the context of the proposal for PPGs led the panel to propose a set of general principles to guide further performance partnership efforts. First, those who are affected by decisions resulting from the application of performance measures (e.g., state and local health agencies) must share fully in the creation and selection of the measures and, where possible, the measurement process. Second, state- and local-level data are essential for achieving federal goals for performance measurement, and ensuring adequate support for data collection mechanisms to produce those data is important to the success of the system. Third, to make efficient and effective use of resources, performance measurement should, to the extent possible, rely on existing data systems that serve other managerial and operational purposes as well. Collaboration across and within the federal, state, and local levels can reduce the inconsistencies and incompatibilities that can arise in independently developed monitoring programs or data systems while preserving sufficient flexibility to accommodate the differences among states and communities in their programs and goals. Finally, performance measurement should focus on monitoring progress toward the overall goals of an activity and not on measuring the impact of a particular source of funding for that activity. The health outcomes that are linked to program goals are influenced by many factors, making inferences concerning the effect of individual funding sources problematic.

A NATIONAL INFORMATION NETWORK FOR HEALTH-RELATED DATA

Further advances in performance measurement for publicly funded health programs will require thoughtful and continuing attention to a varied set of policy, programmatic, and data and information system issues. The current focus on performance-based accountability has helped highlight the limitations of existing health-related data and data sources. Despite their diversity and breadth, the data presently suitable for performance measurement are limited. The need for stronger and more coherent health data systems is a concern at the federal, state, and local levels and in the private sector as well.

DHHS has an important leadership role to play, but all of the participants must share responsibility for ensuring that health data and data systems receive adequate support to operate efficiently and effectively. An investment must be made in the data collection programs and information technology that are at the core of these information systems, as well as in training and technical assistance for the people who produce and use health data.

The panel's conclusion is that over the long term, performance measurement will be easier and more effective with the development of a broadly based national health information network that can promote a coordinated and collaborative approach to meeting diverse needs for a variety of health-related information, including performance data. This does not necessarily require the creation of an entirely new data system or a federally managed system. Instead, the panel envisions building on existing data systems across the local, state, and federal levels to produce the information required for performance measurement.

The development of a truly comprehensive health information network of national scope is a substantial endeavor that will almost certainly require an incremental approach with financial investment from a variety of sources, as well as commitment and persistence. The challenge is to develop a reasonably efficient and effective network that accommodates the many different data sources and information requirements that currently exist. The resulting network should meet real managerial and accountability needs; reflect important interdependencies and relationships across governmental and programmatic lines; and recognize that multiple, specialized data systems may no longer be affordable or consistent with other critical priorities.

As envisioned by the panel, the proposed national health information network should have the following features:

National Collaboration Collaboration among local, state, federal, and private efforts must be supported by adequate resources and must recognize the essential needs and contributions of each participant. The aim is to achieve a network of national scope that is not specifically a federal enterprise.

Linkage but not Consolidation The proposed network does not depend on

the development of a single national health database or information system. Compatible structure and architecture in information systems will promote linkage of comparable data and sharing of information. The confidentiality and security of personal health information must be ensured.

Ability to Meet the Needs of Varied Users The network should support multiple purposes, including monitoring for performance-based accountability of population-based and personal health services, operation of health programs, and delivery and management of clinical care. It should serve the information needs of managers, planners, health care providers, evaluators, policy makers, and the public at the national, state, and local levels, including ensuring that states and communities have the opportunity to produce or obtain timely data of local interest and of sufficient detail to be representative of their populations.

Standardization of Data and Measures Clear and common definitions of data elements, measures, and coding systems and standard approaches to data collection must be established so that information can be aggregated across multiple populations and regions, and comparisons, where appropriate, can be made over time or among populations and geographic areas.

Appropriate Performance Measures The measures adopted should represent multiple perspectives, specifically including that of the consumer (e.g., satisfaction with access to services). They should be applicable to varied types of services (e.g., individual and population-based), to different age groups, and to people of differing ethnic and cultural backgrounds.

Efficient and Effective Use of Resources Coordination and integration of information systems will often prove more efficient than the development and operation of program-specific systems. Data collection should be guided by careful judgments about the intensity of the information required for performance measurement in terms of frequency of data collection, level of detail, and completeness of coverage (e.g., sample or census data). The network should provide access to appropriate information from sources that are not primarily health related (e.g., highway safety, corrections). Staff at all levels must have sufficient training and technical assistance to manage data systems and use information effectively.

Adaptability to Change Performance measurement and a national health information network must be able to change and evolve as information sources grow, knowledge expands, or program priorities and activities change. New measures must be developed and tested to respond to these evolving needs.

The panel concluded that further progress in performance measurement and in the development of a multilevel, user-oriented national health information network requires work in four broad areas: (1) policy actions that promote collaboration by federal, state, and local stakeholders in the performance measurement process and more effective integration of data systems; (2) operational principles that support efficient use of resources and promote use of appropriate perfor-

mance measures and data; (3) essential investment in data systems and in the staff who collect and use performance information; and (4) research to improve the evidence base for performance measures, as well as the use of performance measurement. The panel recommends several specific steps in each of these areas.

POLICY ACTIONS

National Collaboration

To succeed, performance measurement and the national health information network proposed by the panel must be supported by a broad and continuing commitment to a collaborative process that brings together as partners stakeholders from across federal, state, and local governments. The panel advocates such a national collaborative partnership to ensure that performance measurement practices and health data systems are responsive to program priorities and information needs at all levels of government. In addition to *intergovernmental* collaboration, the panel advocates greater *intragovernmental* collaboration to reduce duplication of effort and to promote data sharing and the development of comparable measures and definitions for data related to health outcomes and program activities. Where appropriate, these public-sector efforts should interact with related activities in the private sector.

The national collaboration recommended by the panel will require a process for initiating and continuing consensus-building discussions. DHHS is a key participant and may be an essential catalyst for this process, but it must act as a partner with state and local stakeholders. To ensure full and fair consideration of multiple points of view, participants may wish to identify an interested party without a direct stake in the outcomes (e.g., a foundation, a university) that can convene local, state, and national stakeholders in a neutral setting. Stakeholder groups may also wish to establish well-defined mechanisms for designating their representatives in these discussions. Participants should include both staff with policy and programmatic responsibilities who use health data and those with technical expertise in data collection and data analysis who produce and manage health data.

A well-designed and effectively operating performance monitoring system promises benefits for all of its participants. Those participants must, however, share responsibility for the design and maintenance of the system and for investment of the resources needed to give the system sufficient capacity to operate effectively. Participants' responsibilities also include working toward compromise solutions in such matters as uniformity in definitions and procedures, the choice of data items and data collection methods, and the timeliness and format of data forwarded to other collaborators.

To establish a collaborative base for performance measurement and a health information network, the panel recommends the following steps:

- 1. Federal, state, and local governments should commit to a common and national strategic goal of incorporating performance measurement into the practices of publicly funded health programs.**
- 2. Federal, state, and local governments, with input from private partner organizations, should plan and implement all steps of the performance measurement process in full collaboration with one another.**
- 3. DHHS should work in partnership with members of the relevant groups representing policy, program, and technical officials of states and local entities to establish a process for developing policies and procedures that can facilitate the implementation of performance measurement efforts in health-related areas.**
- 4. Federal, state, and local governments should accept explicit responsibilities, determined in collaboration with other stakeholders, in return for their share in the governance of and benefits from broader efforts to improve performance monitoring.**

Integration of Data Systems

The categorical nature of much of the federal funding for state and local health-related programs has often encouraged both a fragmented approach to health problems and the development of program-specific data systems and reporting requirements. A strictly programmatic perspective may discourage a more comprehensive approach that can capitalize on the complementary, overlapping, and even synergistic interactions among programs and their information system needs.

Even though programmatic funding streams are likely to remain a prominent feature of federal funding, additional opportunities are needed to improve data systems at the state and local levels by coordinating and integrating a broader array of health data. Some federal agencies are supporting a more integrated approach, and the panel encourages other agencies to facilitate a broader perspective in the planning for information system changes and to improve the likelihood of generating additional funds for the implementation of those changes. Specifically, the panel makes the following recommendations:

- 5. DHHS should lead efforts to integrate data systems across categorical health program lines.**
- 6. DHHS, in collaboration with state and local partners, should review restrictions on the use of grant funds to determine whether**

they represent a significant barrier to progress in the development of integrated health information systems. If so, DHHS should pursue changes in the terms of those grant programs that would permit greater flexibility in the use of the funds.

Technology Policy

Rapidly evolving information and communications technologies will enhance the feasibility of performance measurement and the health information network envisioned by the panel. Many state and local health agencies, however, lack the resources to keep up with the rapid pace of the revolution in information technology. Effective use of these technologies requires the development and implementation of standards to facilitate the transmission, aggregation, and linkage of data from multiple sources without requiring the standardization of equipment or operating systems and software. DHHS should serve as a catalyst for consensus building on information collection and transmission standards, and as a resource for technical assistance in the application of new information technologies for a broad range of health-related data. The panel emphasizes that development and use of information technologies must always address protection of the confidentiality and security of health-related data. Specifically, the panel makes the following recommendation:

7. DHHS should provide leadership in the development and use of data transmission standards and of new information technologies to collect, analyze, and disseminate health-related data.

OPERATIONAL PRINCIPLES

The panel identified several matters related to the availability of data and the further development of performance measures that should be addressed to promote the successful implementation and operation of a performance measurement system for publicly funded health programs.

The development of a performance monitoring system requires consideration of the broad range of factors that influence desired health outcomes, as well as the administrative, analytic, and technical resources needed to collect data and use performance measures. State and local government agencies whose responsibilities are not primarily health related must be part of the process. This includes agencies with programmatic responsibilities in nonhealth areas (e.g., criminal justice, housing, transportation) and those that collect and manage data on basic socioeconomic characteristics of the population and the state (e.g., population estimates, economic development data).

Although few existing data systems have been designed specifically for performance measurement, they nonetheless provide an essential base from which to

build an information network that can meet a broad range of health information needs, including performance data. The panel recommends that, instead of creating an entirely new data system for performance measurement, data needs be met to the extent possible by using existing systems, such as vital records, notifiable disease systems, adult and youth behavioral risk factor surveillance, cancer registries, and records on client services (e.g., prenatal care, substance abuse treatment). Other data systems beyond the purview of health agencies can also provide information valuable for tracking health risks and outcomes.

A performance monitoring system requires not only continuing data collection activities, but also a mechanism for ongoing review and refinement of performance measures. Measures must change to reflect the evolving knowledge base on which they rest, changes in health needs and opportunities for intervention, and changes in the health policy environment. A broad range of stakeholders must participate in the review process to ensure that performance measures are consistent with state and local public health priorities. Policy, programmatic, and technical perspectives must all be considered. The collaboration between the Centers for Disease Control and Prevention and the Council of State and Territorial Epidemiologists for periodic review and revision of the case definitions of specific infectious diseases might serve as a model in this regard.

As the panel observed in its first report, data limitations, differing health problems, and differing program priorities preclude prescribing a single set of performance measures for use by all states and communities. Instead, DHHS should work with states and communities to assemble sets of measures that meet the basic tests of validity, reliability, responsiveness, and data adequacy; users can then select smaller subsets of measures that meet their specific needs. Because states and localities may reasonably pursue many different strategies to target a single health outcome, a large number of process and capacity measures should be available for user choice.

Standard definitions for performance measures and standards for data collection should be adopted to enhance the comparability of performance data over time and across states and localities. Use of common measures and data definitions may encourage cost efficiencies by reducing the need to redesign data collection instruments, electronic processing protocols, and similar infrastructure elements. With sufficient comparability across state data systems, greater reliance might be placed on aggregating state data to produce national measures, rather than requiring separate data collection systems at the federal and state levels. Stakeholders must have a means of achieving consensus on these standards and harmonizing the implementation of their performance measurement activities. The advantages of greater standardization should not, however, obscure the need for continued critical assessment of the appropriateness of the measures and methods being used.

It is essential to ensure that performance measurement rests on data and data analysis of high quality. Moreover, differences in data quality across individual

information systems may reduce the comparability of performance data. Although no data system or data set is ever perfect, and costs to reduce residual errors can be high, quality standards must be adopted. These standards should be based on informed assessments of how the data are to be used and the degree of accuracy and precision required to serve those uses.

The panel makes the following recommendations regarding these issues:

8. As states and communities work to implement performance monitoring systems for health-related programs, they should ensure that all relevant public agencies, including those outside traditional health areas, have the opportunity to participate.

9. When possible, partners should obtain performance measurement information from existing or enhanced federal, state, and local information systems.

10. DHHS, in partnership with state and local stakeholders, should lead the implementation of a process for ongoing development and review of performance measures to be used in conjunction with state and local health programs.

11. DHHS, in partnership with state and local stakeholders, should lead a process for assembling and evaluating sets of performance measures from which users can identify and agree upon those appropriate for specific applications.

12. DHHS should work in partnership with state and local stakeholders to promote the development and adoption of standard definitions for performance measures and standards for associated data collection and data quality in performance measurement systems.

ESSENTIAL INVESTMENT

Performance measurement activities are likely to impose new demands on those whose performance is being assessed (e.g., additional data collection or data system development, new data analysis and reporting). Adequate resources must be made available to meet those demands, as well as to maintain the effective elements of current data systems. Furthermore, efforts to enhance a health information network should not compromise funding for program services. To respond to these concerns, the panel recommends investments both in data systems and in training and technical assistance.

Data Systems

Existing health data systems provide a strong base for performance measurement, but they generally have not been developed for this purpose. Adequate resources are needed to maintain key information systems and to enhance or develop new systems for performance data that cannot currently be produced. Investments in data for state- and local-level performance measures should be a high priority. Innovative ways to use these resources should be explored, as should opportunities to improve the current investment of resources in data collection and analysis activities at the federal, state, and local levels. For example, careful examination of duplication in current data collection and data systems may suggest more efficient ways of meeting information needs, potentially freeing resources to improve or expand data systems. To address these matters, the panel makes the following recommendations:

13. DHHS and state and local users of performance measurement data should each commit resources to reduce gaps in the supporting information systems.

14. DHHS should sponsor a review of the current array of federal, state, and local data collection and analysis activities to begin an assessment of how existing resources might be used most effectively to meet performance measurement and other needs for health data. This review must include participation by appropriate state and local representatives.

Training and Technical Assistance

An investment is also necessary in state and local capacity for data collection and analysis. Staff vary in their knowledge of the relevant disciplines and methodologies (e.g., epidemiology, statistics, social science research) and in their experience with the use of data to plan, evaluate, and revise community programs. Expertise is also required in such areas as hardware, software, systems design and integration, and applications development. State and local health agencies are often understaffed, making it difficult for them to assume additional tasks associated with performance measurement. Relatively low salaries also place most health agencies at a disadvantage in the current highly competitive information technology market. If DHHS were to support a central resource for information and guidance on technology matters, state and local health agencies might be able to make more rapid use of a broader range of expertise than they could assemble on their own. Therefore, the panel recommends the following steps:

15. To ensure the success of performance measurement, all stakeholders, with substantial leadership from DHHS, should contribute ongoing technical assistance, training, and resources to enhance state and local data systems and analytic capacity.

16. DHHS should develop and maintain information technology expertise to assist states and communities as they use new technologies to improve the quality of and capacity for data collection, analysis, and dissemination.

A NATIONAL RESEARCH AGENDA

Research must be an integral part of any ongoing program of performance measurement for health-related programs. Because experience with performance measurement is still limited, studies are needed to improve understanding of what measures and methods of data collection are appropriate. Further research must also be done to establish the evidence base for causal links between program interventions and desired outcomes. This evidence, essential for selecting demonstrably meaningful capacity, process, and risk status measures, is currently limited in many fields. Studies will need to draw on expertise from a variety of disciplines, and they must be informative for a variety of settings at the local, state, and national levels. However, research by itself is not sufficient for informing and improving the performance monitoring process; resources must also be available to ensure that significant findings are communicated to those involved in performance measurement. The panel makes the following specific recommendations:

17. Federal agencies, foundations, and other private-sector groups should develop and fund a research agenda to support performance measurement activities, including the testing of intervention effectiveness, the investigation of the links between program capacity and processes and program outcomes, the development of measures, the refinement of data collection and information system technologies, and the use of performance measurement systems and performance-based decision making.

18. DHHS, foundations and other private organizations, and other partners involved in performance measurement activities should contribute in an appropriate manner to a process of information gathering and dissemination to support the use of evidence-based performance measures.

FINAL OBSERVATIONS

The broad national health information network envisioned by the panel should enhance the capacity of federal, state, and local health programs to meet performance measurement obligations and to use performance data and other information to achieve desired health outcomes. The leadership of DHHS and its agencies is critical in this process, but the department must participate as a partner with stakeholders at the state and local levels, in the public and private sectors, and across a variety of program areas. Several significant challenges lie ahead. Appropriate performance measures must be developed, and the data needed to use those measures must be available. Greater consensus must be achieved regarding standards for measures and data that will promote comparability in performance measurement. Necessary technical and analytic skills must be developed and applied to the creation and use of performance data. Sufficient financial and nonfinancial resources must be obtained to support both near-term efforts to introduce performance measurement activities and longer-term data collection, analysis, and research necessary to sustain those efforts. The need to address these issues is great, and the current commitment to performance measurement presents an opportunity to make significant progress toward meeting this need.

1

Introduction

As in many areas of public and private endeavor, publicly funded programs intended to protect and improve the health of the public are being asked to account in measurable ways for their performance. The 1990s have brought a growing emphasis on accountability for achieving desired outcomes, and methods of performance measurement have emerged as essential tools for operationalizing this quest for accountability. A system of performance measurement promises improved documentation of the contributions of public and private agencies, and can serve as a quality improvement tool by drawing attention to practices shown to contribute to desired outcomes and by identifying areas needing improvement. In fact, many people who are well informed about public health, health policy, health economics, and related matters believe that we cannot expect public funding to increase or even be maintained at current levels without better documentation of the return on program investments.

Measuring performance is not a new idea, but the emphasis on outcomes has changed the way we think about these issues and what needs to be measured. It is no longer enough to ask, “How many people enrolled in a smoking cessation program?” or even “How many people finished the program?” Now, answers are also sought to questions such as “How many people stopped and are still not smoking a year after finishing the program?” Selecting the right questions requires an understanding—still limited in some fields—of the often complex relationships between program activities and health outcomes. Answering the questions requires access to appropriate data. Existing data sources, however, have generally not been created for this purpose and may not be readily adaptable to meet the need.

In 1995, the U.S. Department of Health and Human Services (DHHS) proposed the establishment of Performance Partnership Grants (PPGs) requiring the application of performance measurement methods to a set of federal block grant programs that provide funding to states for public health, substance abuse, and mental health activities. That proposal made it necessary for DHHS to consider what the appropriate performance measures would be, how they would be used, and whether suitable data were or could be made available to support the process. The department sought assistance in addressing these issues from the Committee on National Statistics of the National Research Council. The Panel on Performance Measures and Data for Public Health Performance Partnership Grants was assembled in fall 1995 to assess the state of the art in performance measurement for program areas covered by the specified block grants and to recommend steps toward improving performance measures and performance measurement for health-related programs.

The work of the panel has resulted in two reports, of which this is the second. In its first report (National Research Council, 1997), the panel discussed specific measures that are feasible to use now in connection with the block grant programs, as well as conceptual and policy issues related to the use of performance measures. In this second report, the panel looks beyond measures for specific program areas to address broader data and information system issues that require attention at the federal, state, and local levels to advance the practice of performance measurement for publicly funded health programs.

ORIGINS OF THE STUDY

The immediate impetus for this study was the DHHS proposal to establish PPGs for a specific set of health programs. That PPG proposal was, however, a reflection of a more general interest in performance measurement that is evidenced by parallel developments in public health, health care, and public policy. All of these developments have helped draw attention to the challenges of identifying appropriate measures and obtaining high-quality data.

Performance Partnership Grants

States receive DHHS grant funds in support of various health programs. Seeking a way to increase state flexibility in the use of these funds while enhancing accountability for progress toward program goals, DHHS proposed that formal legislative changes be made for some of these grant programs to mandate the implementation of PPG arrangements between states and the federal government. The program areas covered by the original PPG proposal were chronic diseases; sexually transmitted diseases (STDs), human immunodeficiency virus (HIV) infection, and tuberculosis; immunization; mental health; substance abuse; and

three areas of special interest to DHHS—sexual assault, disabilities, and emergency medical services.

The proposal called for DHHS and each state to negotiate an agreement on program objectives for a 3- to 5-year period. Each agreement would also include a set of related performance measures to be used as a basis for monitoring progress toward those objectives. The PPG concept envisioned that DHHS, in consultation with states, public health professionals, private organizations, public agencies, and citizens, would develop a menu of performance measures from which states would select a subset appropriate to their program goals. Because the problems and priorities of states vary, a single set of required measures for use by all states was not considered appropriate.

It was originally expected that the PPG mechanism would be formalized through legislation, but this has not happened. Nevertheless, the idea of performance partnerships based on negotiated federal-state agreements regarding program objectives and measures remains viable and is being implemented for certain grant programs (e.g., the Maternal and Child Health Block Grant; see Maternal and Child Health Bureau, 1997).

Other Influences

Perhaps the most direct antecedent to the PPG proposal is the Government Performance and Results Act (GPRA) of 1993, which requires the federal government to measure the performance of all federal programs. This requirement has focused the attention of DHHS agencies on the issue of performance measurement and gives them an incentive to implement performance reporting for their grantees.

With its emphasis on the collection and analysis of data related to outcomes, performance measurement has close ties to the systematic assessment of health status and health needs that is recognized as a core function of public health (Institute of Medicine, 1988). At the federal level, these assessment activities already encompass the compilation and publication of vital statistics and disease surveillance data collected by states; survey programs such as the National Health Interview Survey; and ongoing public health monitoring efforts, including those to track progress toward the health promotion and disease prevention objectives of Healthy People 2000 (soon to be updated by Healthy People 2010). Similar activities are conducted by state and local health agencies.

The health care field has responded to concerns about assessing and improving the quality of care with a variety of performance measurement activities. An early emphasis on quality assurance encouraged a focus on finding and responding to errors. There has been a gradual shift to a quality improvement approach that puts greater emphasis on using measurement to monitor processes and guide their improvement so better health outcomes can be achieved. The proliferation of new models of health care organization and delivery has led to further changes,

including the development of sets of standard measures for the processes and, increasingly, the outcomes of care (e.g., the Health Plan Employer Data and Information Set [HEDIS]; see National Committee for Quality Assurance, 1997). These efforts have been conducted under the aegis of independent organizations such as the National Committee for Quality Assurance, the Foundation for Accountability (1998), and the Joint Commission on Accreditation of Healthcare Organizations (1998b).

CHARGE TO THE PANEL

For this study, the panel was charged with the following tasks: (1) identify measurable objectives that states and other interested parties might want to achieve through PPG agreements, and that can be monitored at the state and national levels either now or with small modifications to existing data systems; (2) identify measures relevant to PPG agreements that cannot be assessed, but are important to states and the federal government and therefore require further development; and (3) recommend improvements to state and federal surveys and data systems to facilitate the future collection of data for both existing and developmental measures.

The panel's first report (National Research Council, 1997) focused primarily on task 1. The panel addressed broad analytic and infrastructure issues involved in developing and using performance measures. In addition, the panel assessed more than 3,200 candidate PPG measures, proposed by more than 1,500 participants at four regional meetings and by professional associations. Some 60 health outcome and risk status measures were selected as representative of those that might be used in conjunction with federal-state PPGs in the program areas covered by the original PPG proposal. (See Appendix A for a list of the health outcome and risk status measures proposed in the panel's first report.) Related process and capacity measures were also suggested. (The various types of measures are defined in the next section.) In the areas of mental health and substance abuse, a lack of consensus on outcome measures and the limited availability of comparable data collection across states led the panel to frame that portion of its discussion in terms of measures that might be used, but would require further development (i.e., an approach more consistent with task 2). The major findings presented in the panel's first report are reviewed briefly later in this chapter.

The present report takes up tasks 2 and 3. The panel has given somewhat less attention to the identification of additional measures (task 2) than to improvements in data collection and information systems (task 3) for several reasons. First, the broader issues related to the development of performance measures and data systems that are addressed by task 3 were identified in the first phase of the study as being of higher priority and requiring more immediate attention than the narrower concerns of task 2. The panel felt further that since federal-state performance partnership agreements are now expected to develop in program areas not

included under the original proposal (e.g., maternal and child health), consideration of measures needing further development should not be restricted to the original programs. Likewise, the panel concluded that it is important to rethink the program-specific perspective of the PPG proposal as a basis for conceptualizing measures of health outcomes or enhancing data systems. The panel also recognized that the discussion could and should be broadened beyond the federal-state PPG framework to include the local level as well.

ESSENTIAL DEFINITIONS

Performance Measurement and Related Concepts

The term “performance measurement” is used in various contexts. In this report, the term denotes the selection and use of quantitative measures of program capacities, processes, and outcomes (assumed to be health outcomes in this case) to inform the public or a designated public agency about critical aspects of a program, including its effects on the public. The related term “performance monitoring” is used here in the context of a continuing set of performance measurement activities. A “performance measure” is the specific quantitative representation of a capacity, process, or outcome deemed relevant to the assessment of program performance.

One of the principal purposes of performance measurement is to assess whether progress is being made toward desired goals and whether appropriate program activities are being undertaken to promote the achievement of those goals. Performance measurement can also serve to identify problem areas that may require additional attention or, more positively, successful efforts that might serve as models for others. In some fields, performance measurement is being used under certain circumstances as a tool for regulation and resource allocation. The panel has advised against the use of performance measures for resource allocation for health programs until an adequate understanding is developed of the causal relationships between program activities and outcomes, of the measures and data needed to represent those relationships adequately, and of the appropriate adjustment methods for comparisons of dissimilar populations. Well-designed research and evaluation studies are needed to reveal more about the causal relationships that may exist between program activities and outcomes. Even with such studies, the panel cautions against using performance measures as the sole basis for causal inferences regarding program performance because of the diversity of factors beyond program activities that affect most health outcomes. See Chapter 2 for additional discussion of the characteristics and uses of performance measurement.

“Accountability” for performance—an obligation or willingness to be assessed on the basis of appropriate measures of actions and outcomes with regard to the achievement of program or policy purposes—is an essential element

of the results-oriented management approach within which performance measurement is usually applied. Accountability can be required of government units through legislative or executive mandate. With GPRA, for example, the Congress has created a requirement that the executive branch agencies develop performance plans with appropriate performance measures. Under the kinds of performance partnership agreements represented by the PPG proposal, however, states incur an obligation to report on performance by accepting federal grant funding, but they are recognized as partners with whom some of the terms of an agreement are negotiated rather than dictated. In some cases, causal relationships between program activities and outcomes may be clear enough to justify holding the program directly accountable for observed outcomes. More often, and especially for complex matters such as health and well-being, requirements for accountability cannot be translated into an assumption that accountable parties always bear sole responsibility for the outcomes they report (Wholey and Hatry, 1992). In either event, continued failure to make progress toward intended performance goals should trigger analysis and change in policy and programs.

Performance measurement is also a prominent aspect of efforts to assess the quality of health care (see, e.g., National Committee for Quality Assurance, 1997; Foundation for Accountability, 1998; Joint Commission on Accreditation of Healthcare Organizations, 1998a,b). However, the focus on quality of care differs in some respects from the panel's objective of measurement and reporting for the purpose of monitoring program performance. GPRA and the PPG efforts are specifically tied to government activities, but private-sector and provider-led organizations are playing a substantial role in clinical quality assessment. Measurement primarily for internal quality assurance and quality improvement purposes has been supplemented by the development of measures and external reporting programs to help employers and other purchasers of health services, as well as regulators and policy makers, compare the performance of provider groups. Measures and reporting formats that can be useful to individual consumers are also being studied.

Categories of Performance Measures

In its first report, the panel emphasized the need for several types of measures to assess program performance: health outcome, risk status, process, and capacity (see Box 1-1 for the definitions used in that report). Some health outcomes of primary interest, such as reductions in mortality or morbidity, may be impractical to measure as indicators of program performance. The time lag between an intervention and changes in those outcomes is too great for the effects to be observable within the relatively short time frames (e.g., ranging from 3 to 5 years in the PPG proposal) used to monitor program performance. To provide a partial solution to the problem posed by long latency periods, the panel included measures of risk status as intermediate outcomes. For a risk status measure to be

Box 1-1
Categories of Performance Measures

Health Outcome: Change (or lack of change) in the health of a defined population related to an intervention, characterized in the following ways:

- health status outcome:* change (or lack of) in physical or mental status
- social functioning:* change (or lack of) in the ability of an individual to function in society
- consumer satisfaction:* response of an individual to services received from health provider or program

Risk Status (intermediate outcome): Change (or lack of) in the risk demonstrated or assumed to be associated with health status.

Process: What is done to, for, with, or by defined individuals or groups as part of the delivery of services, such as performing a test or procedure or offering an educational service.

Capacity: The ability to provide specific services, such as clinical screening and disease surveillance, made possible by the maintenance of the basic infrastructure of the public health system, as well as by specific program resources.

SOURCE: National Research Council, 1997:9.

appropriate, of course, there should be consensus that the result being measured is directly related to the health outcome of interest, although it is rarely possible to account adequately for all of the many confounding factors that affect the ultimate health outcome. Similarly, process and capacity measures should have a recognized and generally accepted relationship to relevant health outcomes.

For example, a state with a goal of reducing its mortality rate from breast cancer could seek to reduce the risk of death by increasing the number of mammograms provided to women aged 50 and over. The mammography rate could then be used as a risk status measure. In addition, the state could track changes in processes (e.g., health education programs, requirements that private insurers include coverage of specific activities such as mammography or surgical treatment, and postoperative follow-up care) and elements of capacity (e.g., numbers of trained staff and facilities offering mammography screening) that are

believed to be related to the level of mortality from breast cancer. A detailed set of such measures could provide some understanding of the particular services that are available and that may be contributing to or inhibiting desired changes.

PHASE I: FOCUS ON SELECTION OF PERFORMANCE MEASURES

In its first report, the panel identified various outcome, process, and capacity measures that it considered suitable for federal-state performance partnership agreements under the specific grant programs for which PPGs had been proposed. The panel emphasized that these particular measures were representative examples, not a definitive or exhaustive list. Because health needs and program priorities, as well as data resources, vary among states and will surely vary over time, all of these measures will not be appropriate for every state and every future need. This is especially true for the process and capacity measures. States can pursue many reasonable strategies to improve health outcomes, and each strategy may require a different set of process and capacity measures. To illustrate the range of potential strategies and the implications for process measures for a single program goal, Table 1-1 (reprinted from the panel's first report) lists examples of strategies for reducing the incidence of tobacco smoking and process measures associated with each strategy.

As part of phase I of the study, the panel also addressed broader issues of performance measurement by providing a general analytic framework for use by states and DHHS in assessing the appropriateness of outcome, process, and capacity measures for individual performance agreements. Recognizing that data resources and measurement methods need improvement, the panel recommended in its first report that DHHS continue to work with states and local areas toward several infrastructure goals: developing common definitions and measurement methods; encouraging efficient development of data resources that would support multiple public health, mental health, and substance abuse needs; incorporating state and local data priorities in national infrastructure development efforts; and promoting state and local data collection and analytic capabilities. These issues are addressed more thoroughly in the present report.

The principal conclusions and recommendations of the panel's first report are briefly reviewed below.

Use of Measures of Process and Capacity as Well as Outcomes

Despite their widespread use and intuitive appeal, health outcome measures by themselves are insufficient for monitoring the effectiveness of a given program in achieving health goals. One reason is that outcomes are often influenced by factors other than activities associated with a particular program or agency. An example is mammography rates for women over age 50, which can be affected by factors such as state-sponsored consumer education, private advertising, tech-

TABLE 1-1 Examples of Program Strategies and Related Process Measures for Reducing the Incidence of Tobacco Smoking

Program Strategy	Process Measure
Limit illegal youth purchases of smoking tobacco	Percentage of vendors who illegally sell smoking tobacco to minors Percentage of communities with ordinances and regulations restricting smoking tobacco sales Number of vending machines selling smoking tobacco in locations accessible by youth Presence or absence of state or local tobacco retailer licensing system
Increase the price of tobacco products	Amount of excise tax (cents) per pack of cigarettes
Restrict smoking tobacco advertising	Percentage of communities with ordinances or regulations restricting smoking tobacco advertising Number of billboards advertising smoking tobacco close to schools and playgrounds Number of sport or entertainment events sponsored by tobacco companies
Restrict indoor tobacco smoking	Percentage of worksites (day care centers, schools, restaurants, public places) that are smoke free (have limited smoking to separately ventilated areas)
Educate children about hazards of smoking tobacco	Proportion of elementary, junior high, and high schools with age-appropriate smoking prevention activities and comprehensive curricula
Increase access to or availability of smoking cessation programs	Proportion of current tobacco smokers visiting a health care provider during the past 12 months who received advice to quit Proportion of managed care organizations (or schools or obstetric and gynecological service providers) that have active smoking prevention and cessation plans
Market effective antismoking messages to the general public	Percentage of adults who can recall seeing an antismoking message during the 12 months following a media campaign

SOURCE: National Research Council (1997:20).

nological changes that affect cost, and changes in insurance coverage. For substance abuse and mental disorders, knowledge regarding the factors that influence the longer-term outcomes of these chronic and recurring conditions is particularly limited. A second important limitation on the sole use of outcome measures to monitor program effectiveness, noted earlier, is the impractical delay involved in observing certain outcomes of interest, such as the length of time required for many cancers to become detectable. A third limitation is the rarity of some important outcomes, such as major outbreaks of food- or water-borne illness. Relying only on the detection of an outbreak of cryptosporidiosis, for example, would not be an acceptable means of monitoring the effectiveness of water treatment services.

The panel therefore concluded that performance monitoring must also make use of measures of intermediate outcomes, process, and capacity for which scientific evidence or professional consensus has established a relationship to the desired health outcome. Even this “multimeasure” approach may not provide conclusive evidence of the effectiveness of particular interventions, but it will allow interested parties to examine actions taken by agencies to realize their objectives and consider whether changes in the magnitude or direction of their efforts are needed.

Guidelines for Selecting Performance Measures

The panel applied four guidelines in its review of the proposed PPG measures and urges others to use these same guidelines when selecting performance measures:

1. *Measures should be aimed at a specific objective and be result oriented.* Outcome measures must clearly specify a desired health result, including identifying the population affected and the time frame involved. For process and capacity measures, the link to a health outcome should be clearly specified.

2. *Measures should be meaningful and understandable.* Performance measures must be seen as important to both the general public and policy makers at all levels of government, and they should be stated in specific but nontechnical terms.

3. *Data should be adequate to support the measure.* Data must meet reasonable statistical standards for accuracy and completeness; be available in a timely fashion, at appropriate periodicity, and at reasonable cost; and be collected using similar methods and with a common definition throughout the population of interest. Comparisons across states or other population groups are valid only if definitions and collection methodologies are consistent across those populations.

4. *Measures should be valid, reliable, and responsive.* To be valid, a measure should capture the essence of what it purports to measure. To be reliable, a measure should have a high likelihood of yielding the same results in repeated

trials and therefore low levels of random error in measurement. To be responsive, a performance measure should be able to detect a change. It is also important to recognize that a measure meeting these requirements for one purpose may not meet them for another. For example, the infant mortality rate is usually considered a valid and reliable measure of the change in a state's rate of infant death from one period to another. It may not, however, be a valid measure of the performance of an individual public health agency that has only limited influence on factors affecting infant health. Moreover, it may not be a reliable measure of change at the local level because the small number of infant deaths at that level makes the measure subject to random variation from year to year. And it may not be a responsive measure for assessing the impact of a new prenatal counseling program serving a segment of a community that accounts for only a small share of the community's infant deaths.

Limitations of a Program-Specific Approach to Performance Measurement

For the first phase of this study, the panel was asked to consider performance measures that could be used in connection with federal grants to states for the specific program areas noted earlier (i.e., chronic diseases; STDs, HIV infection, and tuberculosis; mental health; immunization; substance abuse; and three areas of special interest to DHHS—sexual assault, disabilities, and emergency medical services). Clearly, the individual diseases and health conditions that the panel studied are only a subset of those that are of concern around the country. The panel believes, for at least three major reasons, that over the long term it would be preferable to monitor performance using a more comprehensive and less program-specific approach that integrates generic with program-specific measures.

First, the use of performance measures to assess the impact of a particular federal funding program is complicated by the fact that those federal funds are often only one of several sources of support for a state or local health program. For example, the federal mental health block grant represents only about 4 percent of state mental health agency budgets, with state general revenues, private insurance, Medicaid, and local sources making up the balance. Because those federal funds do not buy specific services, it appears unlikely that a change in any statewide measure of mental health outcomes could be attributed unequivocally to a mental health block grant.

Second, a program-specific approach to monitoring performance tends to overlook the synergies that can result from the coordination of efforts supported by separate funding programs. For example, both a maternal and child health program and an STD program might target HIV testing in pregnant women, or resources for chronic disease prevention and environmental health might target lead abatement interventions. Given current levels of knowledge, efforts to

attribute outcomes to one or another partial funding source are expensive, often futile, and of no benefit.

A related consideration is efficiently meeting various programs' overlapping data needs. A strictly program-specific approach might lead to duplication of data collection efforts or missed opportunities to adopt measures that can be used by more than one program. For example, measures related to tobacco use may be of interest not only to a tobacco control program but also to programs aimed at preventing cancer, preventing and controlling chronic respiratory illnesses such as asthma, and reducing the incidence of low-weight births.

Finally, and much more broadly and subtly, the program-specific approach has led to hierarchical concepts about the governance, competence, and focus of performance measures and appropriate data systems that support them. A national perspective and federal leadership remain important, but an effective performance measurement and accountability system also requires that state and local agencies play a greater role in defining program priorities and shaping performance measurement activities. Effective change will require true partnership in this endeavor. These new concepts and their implications are discussed further in the succeeding chapters.

Need to Strengthen State and Local Capacity for Data Collection and Analysis

The panel concluded in its first report that the data infrastructure required to support state- and local-level performance monitoring needs to be strengthened. Many federal data collection programs produce national but not state- or local-level rates. Many of the potential health outcome measures identified by the panel are heavily dependent on a small number of collaborative state-federal surveys, such as those of the Behavioral Risk Factor Surveillance System and the Youth Risk Behavior Surveillance System. Even these survey programs do not cover all states or apply consistent survey methods across states. The panel therefore recommended viewing the use of performance measures to assess the effectiveness of public health, substance abuse, and mental health programs as an ongoing, long-term public administration effort that requires a strong commitment by the federal government to providing technical assistance and infrastructure support to its partners at the state and local levels.

Inadvisability of Using Performance Measures Alone for Resource Allocation Purposes

Although there is considerable value in using performance measurement to enhance the effectiveness and accountability of publicly funded programs, the development and use of performance measures, particularly for comparisons across states, is not yet—and may never be—a precise scientific process. Under-

standing of the relationships between health interventions and outcomes and between individuals' characteristics and health outcomes is still limited. Such knowledge is essential for making appropriate statistical adjustments for socio-demographic and other relevant factors. Moreover, the complexity of the relationships among health outcomes, program interventions, and other factors in the physical and socioeconomic environments may make it difficult to monitor performance in sufficient detail to ensure that resource allocation decisions are based on consideration of all the appropriate causal factors. In practical terms, timely and comparable data are often unavailable. Consequently, the panel warned that using cross-state comparisons of "performance" as the analytic basis for determining financial rewards or penalties for participating agencies is, at present, highly problematic.

PHASE II: DATA AND INFORMATION SYSTEM DEVELOPMENT TO SUPPORT PERFORMANCE MEASUREMENT

For the second phase of the study, which addressed the needs for data and information system development to support performance measurement, the panel adopted a broader perspective than was suggested by the study's initial focus on state-level performance measurement for federally funded programs in specific areas of public health, substance abuse, and mental health. Rather than pursue a strictly technical assessment of program-specific measures, data collection methods, or analytic techniques, the panel judged it important to put performance measurement in a broader data context and to emphasize the commonalities across programs, while still taking note of some special concerns in specific program areas. The study's second phase continued to focus largely on the public sector, but the panel looked beyond the federal-state relationship that defined the PPG proposal to consider a more general notion of performance partnership agreements that can encompass state and local interests as well.

The panel has three aims for the present report: (1) to highlight important technical and policy issues that must be considered in the further development and use of performance measurement for health-related programs; (2) to describe a health information network that would support performance measurement at the national, state, and local levels; and (3) to present a strategy for developing such a network.

A Vision for a National Health Information Network

Certain elements are fundamental to the panel's vision of the kind of information network that should be developed to support health-related performance measurement. A key factor is the development of a national network through an active collaboration among local, state, and federal agencies. A national approach should ensure that information resources, interests, and needs at each of these

levels, as well as in the private sector, are taken into account, while still allowing the aggregation of data in ways useful for larger geographic and administrative units. DHHS and other federal agencies have an important leadership role to play, but they must work in partnership with others who have an interest in such a network. Indeed, the panel envisions a national network of interacting systems, with data and transaction standards supporting the production of performance data that are comparable across sources. The aim is to find a means of supporting information needs for health-related performance measurement within a broader system that serves other operational, managerial, and analytic purposes. A specialized data system for performance measurement is generally not an efficient or cost-effective goal.

Because health needs, program priorities, and resources differ throughout the country and change over time, an information network useful for performance measurement must be adaptable to these differences and changes. Furthermore, because understanding of performance measures and performance measurement is still evolving, an information network must be able to respond as additional empirical evidence is obtained and better methods of data collection are implemented. Finally, any such information network must provide strong protections for the confidentiality and security of data.

The panel's vision for a national health information network is discussed in detail in Chapter 5.

Critical Issues

This report addresses several issues the panel believes to be critical to further advances in performance measurement for health-related programs. In the development of plans for performance measurement, the assessment of data needs, and the enhancement (or redesign) of data systems to support performance measurement, a primary concern is the need for an integrated perspective and effective collaboration. This collaboration involves multiple partners—federal, state, and local agencies, each with multiple stakeholders, as well as program managers, service providers, private nonprofit organizations, and consumers. Also requiring attention are ways to improve the use of existing data and to develop better performance measures. In addition, performance measurement systems will need to address the quality of the data that are collected and used. Information technologies are creating greater opportunities to apply performance measurement, but using those technologies effectively will require attention to data standards. Successful implementation of performance measurement systems will also depend on the availability of training and technical assistance to ensure that skilled staff can apply appropriate policy, programmatic, and technical expertise. Research to improve the science base for and the development and use of performance measures and performance measurement is essential. Another fundamental concern

is determining what resources are needed to support performance measurement activities and ensuring that those resources will be available.

STRUCTURE OF THE REPORT

This report presents the panel’s findings and recommendations regarding data and information systems to support performance measurement for publicly funded health programs. Chapter 2 examines performance-based systems and the uses of performance measures and performance measurement. Chapter 3 considers the characteristics of various health program areas and the implications of those characteristics for performance goals and performance measurement. Chapter 4 explores factors in the current data and information system environment that must be addressed to advance the use of performance measurement. In Chapter 5, the panel outlines its vision of a national health information network that would effectively support performance measurement as well as other objectives, and makes recommendations to further the development and implementation of such a network.

REFERENCES

Foundation for Accountability
1998 About FACCT. <http://www.facct.org/about.html> (April 15, 1998).

Institute of Medicine
1988 *The Future of Public Health*. Committee for the Study of the Future of Public Health. Washington, D.C.: National Academy Press.

Joint Commission on Accreditation of Healthcare Organizations
1998a Nation’s Three Leading Health Care Quality Oversight Bodies to Coordinate Measurement Activities. Press release. May 19, 1998. <http://www.jcaho.org/news/nb.htm> (June 5, 1998).
1998b Oryx Fact Sheet for Health Care Organizations. <http://www.jcaho.org/perfmeas/oryx/sidebar1.htm> (July 24, 1998).

Maternal and Child Health Bureau
1997 *Guidance and Forms for the Title V Application/Annual Report*. Maternal and Child Health Services Title V Block Grant Program. Rockville, Md.: U.S. Department of Health and Human Services, Health Resources and Services Administration.

National Committee for Quality Assurance
1997 *HEDIS 3.0/1998*. Washington, D.C.: National Committee for Quality Assurance.

National Research Council
1997 *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health*. E.B. Perrin and J.J. Koshel, eds. Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics. Washington, D.C.: National Academy Press.

Wholey, J.S., and H.P. Hatry
1992 The case for performance monitoring. *Public Administration Review* 52(6):604-610.

2

Putting Performance Measurement
in Context

The panel views performance measurement as a management and oversight tool intended to improve desired outcomes by focusing attention on quantifiable measures of those outcomes, on changes (or lack of change) in those measures, and on the processes and capacity being applied to achieve the outcomes. The principal aim of this report is to address technical and policy issues associated with the data and information systems needed to provide effective support for performance measurement for publicly funded health-related programs. Consideration of these issues must, however, take into account the broader policy context in which performance measurement is used. It is within this context that performance goals are defined and then translated into performance measures, for which information systems must be able to produce data of the needed scope and quality. This chapter reviews the characteristics and uses of performance-based management and accountability systems, some of their strengths and limitations, and examples of their application in federal and state government settings and in the private sector for health care organizations. It also notes ways in which such mechanisms rely on data that are already available and some of the potential limitations of those data for accurately assessing performance.

**USE OF PERFORMANCE MEASUREMENT
IN ACCOUNTABILITY SYSTEMS**

As noted in Chapter 1, the movement to increase the accountability of organizations and programs for achieving desired outcomes, particularly in the public sector, has led to renewed interest in performance measurement. This approach

to the management of public programs and policies is believed to be superior to other management approaches that are based on micromanagement, process controls, and oversight of resources and activities, and that place little emphasis on results (Osborne and Gaebler, 1992; Wholey and Hatry, 1992).

Performance-Based Accountability Systems

As defined by the National Performance Review (1993), the guiding principle of governmental performance-based accountability systems is the provision of increased flexibility to lower-level units of government, or “partners,” in exchange for increased accountability for results. This increased flexibility may take the form of consolidation of funding streams, elimination of micromanagement, devolution of decision making, or a reduction in bureaucratic paperwork and reporting. Increased accountability for results means that partners focus on outcomes, rather than inputs and processes, as the basic measures of success. Some accountability systems may use such measures to allocate resources or apply incentives to reward desirable outcomes.

Performance-based accountability systems are being established in the public sector between the legislative and executive branches of governments and between levels of government. The Performance Partnership Grants (PPGs) that were proposed by the U.S. Department of Health and Human Services (DHHS) for several of its state block grant programs are an example of a system intended to operate between levels of government. Such arrangements can be established between federal and state, federal and local, or state and local units of government. Even in the absence of formal PPG legislation, performance partnership agreements can be expected to function in this manner.

Public-sector agencies are extending performance-based accountability into their relationships with the private sector through mechanisms such as performance-based contracting for the delivery of services. For example, state substance abuse or mental health agencies often contract with private providers to deliver publicly funded services. With performance-based contracts, those providers can be held responsible for certain overall outcomes among the people they serve. Performance-based accountability can even be extended to interrelationships in a broad community context. The community health improvement process described by the Institute of Medicine (1997) relies on performance measurement to monitor progress toward health improvement goals and ensure accountability of specific segments of the community for the processes and outcomes for which they have accepted responsibility.

These management and accountability arrangements between and within units and levels of government can be viewed as a substitute for the private sector’s market mechanism (see Wholey and Hatry, 1992). In the private sector, it is assumed that in the long run, the discipline of the marketplace will motivate firms to strive for cost-efficiency and maximization of returns to stockholders.

Measures such as profits, rates of return on investments, and market share can be used to assess a unit's success at maximizing efficiency. Furthermore, market forces and signals provide the sorts of feedback managers need to achieve such objectives.

In contrast, the public sector is not governed by the economic forces of competition and profits. Residents who find their state services inadequate or overpriced generally cannot choose (unless they move) to use the services of another state the way a consumer can choose to buy a competing product. To judge the performance of the public sector, stakeholders must rely on other, noneconomic indicators related to human, social, and natural "capital" that must be preserved and invested wisely. There is less consensus on what these indicators should be than on the economic indicators of business performance.

The performance partnership mechanism is one of a much broader class of performance-based systems that have been considered and used in public-sector management over the past 30 to 40 years.¹ These systems include performance-based accountability, performance-based budgeting, performance monitoring, and benchmarking systems. While differing in various ways, all are predicated on a common view that government agencies and organizations need to be more accountable to legislatures, and ultimately to the public, for the resources they receive, and that this accountability should be based on improvements in the dimensions of well-being that such agencies seek to affect.

Problems encountered in earlier efforts to apply performance-based systems offer lessons for current performance-based approaches (Florida Office of Program Policy Analysis and Government Accountability, 1997; U.S. General Accounting Office, 1997c). The extensive information needs of those earlier efforts were not adequately supported by the available record-keeping systems, staff expertise, and computer and information resources. Thus substantial staff time was necessary to meet reporting requirements. Despite this investment of staff time and other resources in producing the required reports, these efforts had little observable impact on funding decisions. The performance-based management approaches used in the past often lacked key leadership support in the executive and legislative branches of government. Furthermore, the analytic character of these approaches made them insensitive to the political aspects of deci-

¹The U.S. General Accounting Office (1997c) has reviewed previous performance-based management efforts. Those efforts can be traced back to the 1950s and program budgeting, which adopted budget categories based on activities and projects rather than classes of expenditure, such as salaries and capital equipment. In the early 1960s, the planning-programming-budgeting system (PPBS) was an attempt to link budgeting with both short- and long-range program plans. Management by objectives (MBO) was adopted in the early 1970s. It called for annual operating plans and targets for program objectives and focused more on productivity assessments than on outcomes. Over the course of the 1970s, MBO was replaced by zero-based budgeting (ZBB), which relied on an annual reexamination of the components of program budgets to assess opportunities for improved management and efficiency.

sion making. If performance measurement is to succeed, it must avoid problems such as these.

Operation of Performance-Based Systems

While various performance-based systems differ in their particulars, there appear to be several key steps in the operation of such systems. These steps are briefly reviewed here.² In Box 2-1, these steps are illustrated by a hypothetical state immunization program and performance measures suggested in the panel's first report.

Step 1: Develop an explicit set of goals and objectives and articulate strategies for achieving them. The first step for a performance-based system is to delineate clearly the goals and objectives of an agency or program. These goals and objectives are often captured in a strategic plan that includes a mission statement and a discussion of how the goals and objectives will be achieved. Furthermore, a strategic plan may outline the resources that will be used to meet these goals and objectives; it may explicitly stipulate the necessary expenditures as well. As noted earlier, one would expect the goals and objectives to focus on outcomes, not process. Such is certainly the case for the two recent federal initiatives in this area—the National Performance Review and the Government Performance and Results Act (GPRA).

A key part of the process of setting goals and developing strategic plans is identifying and involving a program's stakeholders and balancing their potentially competing interests (e.g., reduce costs, increase services, improve quality, replace one activity with another). Much of the recent literature (see, e.g., Wholey and Hatry, 1992; U.S. General Accounting Office, 1996) has emphasized the importance of involving all stakeholders—including policy makers, agency administrators, local program operators, clients, and in some cases members of the public—in the goal setting and planning processes.

In the case of the performance partnership agreements addressed by this panel, granting agencies (e.g., various DHHS agencies) and grantee agencies (e.g., state and local agencies or organizations) may each have their own goals and strategic plans. Negotiated agreements are the mechanism for identifying the particular set of goals and objectives against which grantees' performance will be assessed.

Step 2: Develop and implement strategies for measuring performance. A performance-based system must have a means of assessing progress toward stated goals. This method of assessment is provided by translating program objectives into measures of performance: quantitative or qualitative characterizations of

²See U.S. General Accounting Office (1996, 1997c) for an overview of the structure of performance-based monitoring, accountability, and budgeting systems.

Box 2-1
Steps in the Operation of a Performance-Based
Management System:
Example of a Hypothetical State Immunization Program

Step 1: Develop goals and objectives and strategies for achieving them.

The strategic plan for a state's immunization program might have as a goal reducing vaccine-preventable illness by (1) increasing the age-appropriate immunization rates among children at 2 years of age and (2) increasing influenza immunization rates among adults aged 65 and older. The plan might call for achieving specific levels of immunization coverage at some point in the future (e.g., in 3 years). The strategies for achieving these goals might include enhancing a childhood immunization registry system to generate reminder notices for parents and creating an immunization awareness program to reach older adults.

Step 2: Develop and implement strategies for measuring performance.

The performance of the immunization program might be assessed using *outcome* measures, such as the incidence of measles, rubella, and other vaccine-preventable diseases among children and the incidence of influenza-related deaths among older adults; *risk status* measures, such as the age-appropriate immunization rates among 2-year-old children or the influenza immunization rate among older adults; and *process* measures, such as the proportion of parents with children under age 5 who report receiving an immunization reminder notice and the proportion of older adults living in the community who report having seen information

outcomes to be achieved if those goals are to be realized, processes to be followed in efforts to achieve those outcomes, or capacity available to support those efforts. Although measures based on outcomes are a high priority, a mix of measures will generally be needed to assess the performance of a program from various stakeholder perspectives (e.g., program managers, funders, consumers). For programs that affect outcomes over the long term (e.g., chronic disease prevention) or that guard against possible but rare adverse events (e.g., water treatment), it may be more meaningful to focus on measures that track risk reduction activities and capacity to respond than on outcome measures that would generally show little change in the short term and few differences from program to program. This panel's first report (National Research Council, 1997) provides an extensive discussion of the categories of measures deemed relevant for health-related programs (see Chapter 1 of the present report for a brief review of these categories).

While the process of measuring performance, especially in terms of relevant

on where they could receive an influenza immunization. As a *capacity* measure, the state might use the proportion of children under age 2 that are included in the immunization registry. The data for these measures would be obtained from several sources. Measures of disease incidence might be limited to those diseases that the state has designated as reportable and for which the state health department collects data. Influenza-related deaths would be tabulated by the vital records system. If reasonably complete, an immunization registry could produce data on immunization rates among young children. A survey (e.g., the Behavioral Risk Factor Survey) would probably be the most effective way to obtain data on influenza immunizations among older adults, immunization reminders received by parents, and awareness of immunization services.

Step 3: Use performance information to improve management practices or resource allocation.

Persistently low or decreasing immunization rates would be a signal to examine the operation of the immunization program more closely. The process and capacity measures selected in Step 2 might reveal program weaknesses that could be remedied, such as improving the completeness of an immunization registry's coverage of young children in the state. Finding that the selected process and capacity measures were at desired levels would signal the need to examine other factors that might account for poor performance. For example, the year's influenza vaccine might have been less effective than usual because of the emergence of an unanticipated viral strain.

outcomes that should be influenced by program activities, is likely to vary from one agency or context to another, the literature on these systems offers general guidance (e.g., Wholey, 1983; U.S. General Accounting Office, 1996). In its work on GPRA, for example, the U.S. General Accounting Office (GAO) (1996:24) has noted the importance of establishing "clear hierarchies of performance goals and measures" that reflect the roles and responsibilities at varying program levels, from planning and oversight to grass-roots delivery of services. GAO comments that the performance measures should be tied to program goals and, to the extent possible, demonstrate the results of program actions that are directed toward achieving those goals. At the broadest policy and management levels, a limited set of measures that focus on key outcomes and actions should be used. Including too many measures at this level can divert attention from key outcomes without improving the usefulness of the performance information as a management tool. These measures must, however, be chosen carefully, espe-

cially if they are to be used to monitor a diverse set of activities, such as those likely to be encompassed by federal block grants to states, since activities that are represented in the set of measures are likely to be seen as having a higher priority than those not represented. A greater number of measures may be appropriate at the more detailed operational levels, such as within a state or community program. Although the specific measures are likely to differ across organizational or operational levels, they should be related to each other through their relationship to activities that contribute to the achievement of program goals.

Once measures have been selected, the necessary data must be collected and used to calculate those measures. For some measures, it may be possible to rely on existing data sources, while other measures may require new data collection or data processing procedures. Meaningful interpretation of performance results may also require data on other factors not directly related to program activities or goals but that can affect the environment in which a program is operating, such as widespread disease outbreaks (e.g., epidemic levels of influenza), natural disasters, or changes in the local economy (e.g., increased unemployment because of layoffs). The completeness, accuracy, consistency, and timeliness of the data must be assessed, but such assessments must be made in light of the tradeoff between the benefits of improving the quality of the data and the cost of doing so. Issues related to producing performance data are at the heart of this report and are addressed at greater length in subsequent chapters.

Step 3: Use performance information to improve management practices or resource allocation. The next step for a performance-based system is to apply the information obtained from performance measurement to assess progress toward desired outcomes. If progress is not adequate, performance information can inform steps taken to improve the likelihood of achieving outcome goals in the future. Some policy makers would like to use performance measures to determine resource allocation, directing additional resources to activities demonstrating “good” performance or reducing resources to those demonstrating “poor” performance. As discussed earlier, however, the panel cautions that use of performance measures in an arbitrary, formulaic approach to resource allocation generally is not appropriate because few performance measures can adequately and unambiguously represent the complex mix of factors that determine outcomes. Only if the measures are based on a definitive causal relationship between capacity and process and the outcome of interest, and if experience has demonstrated that they do not stimulate adverse unintended consequences, might it be reasonable to consider using them as a direct determinant of resource allocation decisions.

The element of accountability that is central to such systems implies that performance data should be reported in a form that is accessible and useful to a program’s stakeholders. It is critical to recognize that performance measurement is not an end in itself; it is a tool that should be used in a continuing process of assessment and improvement.

Applications of Performance Measurement

Information obtained from efforts to measure performance can be used to various ends. This section highlights four potential ways such information might be used, particularly in the context of publicly funded health-related programs. The first two reflect a monitoring and reporting function for a performance-based accountability system. Accountability comes somewhat indirectly through the reactions of administrators and constituents in response to information on how an organization is performing. The latter two applications involve the use of performance information to influence program management and resource allocation more directly. These four applications of performance measurement information are not mutually exclusive, but they do differ in their implications for those whose performance is being measured.

1. *Inform various stakeholders (e.g., administrators, public officials, and citizens) of progress toward stated program goals.* Performance measurement information can be used to compare actual performance with performance targets. Performance data can also be used to monitor progress over time or to compare the progress of multiple groups toward agreed-upon goals and objectives. For such comparisons to be appropriate and meaningful, the performance measurement information must be generated in ways that produce comparable data. For example, a state legislature might want to compare the state's immunization rates for 2-year-olds with the national target of 90 percent that was established in *Healthy People 2000* (U.S. Department of Health and Human Services, 1991). The state might also want to assess progress toward this goal by local immunization programs across the state. Consumer-oriented reporting of performance information is illustrated by "report cards" on health care provider performance, such as that developed in conjunction with the Mental Health Statistics Improvement Program (MHSIP) (MHSIP Taskforce on a Consumer-Oriented Mental Health Report Card, 1996).

2. *Assess program effectiveness.* Performance measurement can contribute to program management and accountability by serving as a primary method of surveillance for program effectiveness. It provides a framework to guide the systematic collection of information on desired outcomes and on the program activities that are specifically expected to contribute to the achievement of those outcomes. This performance information can provide an indication of how well programs are working. In addition, an ongoing performance measurement system can often provide data for assessing the effect of changes in other factors or programs related to health services (e.g., the growth of managed care).

This panel's first report (National Research Council, 1997) advised that health-related performance measurement must include a mix of outcome, risk reduction, process, and capacity measures. The use of risk reduction measures to represent intermediate outcomes is important because, as noted earlier, many

health outcomes are too far “downstream” from program activities for direct causal linkages to be established or for those outcomes to be observed soon enough to be useful for program management. In general, routine and direct measurement of program processes and outcomes is not part of current practice at the state and local levels.

As early, real-time indicators of program effectiveness, performance measures can signal matters warranting more attention. Additional analysis is then needed, however, to define the elements of a successful innovation or diagnose the source of a problem.

3. *Improve program performance.* By providing sentinel markers of program effectiveness, performance measurement can guide program managers and policy makers in steps designed to improve program performance. Performance measurement can help focus the attention of practitioners, researchers, and policy makers on best practices. Attention to and accountability for processes and intermediate outcomes that are under more direct programmatic control than longer-term outcomes will lead to a much-needed emphasis on defining standards of practice in health program areas. From the external perspective of a funding agency, data showing poor performance may signal a program’s need for increased technical assistance and for guidance in identifying appropriate practices and determining how they can be implemented.

Incentives and sanctions are also used to encourage improved performance, but may prove difficult to use effectively in the public sector (Florida Office of Program Policy Analysis and Government Accountability, 1997). They can range from generally intangible positive (or negative) recognition for progress toward stated goals to specific and quite tangible financial rewards (or penalties) based on measured performance. The aim is to motivate program staff or communities to achieve desired outcomes (e.g., immunization rates, access to services, desired community behaviors) by comparing performance measurement information with targets set for program goals.

As noted earlier, the private sector often relies on the prospect of financial rewards or penalties (e.g., profits, loss of market share) to create an incentive for good performance. For public-sector programs that do not operate in a competitive, market-based environment, financial penalties may only make it more difficult to improve performance. Instead other, nonfinancial tools can be used to improve performance. For example, continued poor performance that can be attributed to program mismanagement may call for penalties in the form of increased oversight, reduced flexibility, and more directive program management by the funding agency.

The panel emphasizes that in the abstract, fear of sanctions may be an incentive toward improvement, but the application of sanctions will not, by itself, improve performance. Some observers suggest that fears by staff in state agencies that poor performance results will lead to penalties rather than assistance to improve performance can be a barrier to effective use of performance measure-

ment (Wholey and Hatry, 1992; U.S. General Accounting Office, 1994; Florida Office of Program Policy Analysis and Government Accountability, 1997).

4. *Guide resource allocation and regulation of activities.* Performance measurement information is also being used for allocation of budget resources or as the basis for regulatory control to ensure a minimum acceptable performance. For example, some states have adopted performance-based budgeting systems under which decisions regarding agency budgets are directly linked to measures of agency performance (see below for additional discussion of state systems). The panel suggests that the use of performance measures in this manner for health-related programs is appropriate only when clear standards or substantial experience is available to guide actions in a manner that will avoid unintended adverse consequences. For example, linking funding for substance abuse treatment services to rates of treatment completion might discourage acceptance of clients who appear less likely to remain in treatment.

In general, the panel believes that this process should not be as simple as rewarding or penalizing performance by providing or taking away resources. Indeed, as suggested earlier, such an approach may be counterproductive. Take, for example, a county with low immunization rates that have failed to improve over time. This situation could be the result of program mismanagement and poor decision making, or it could reflect especially intractable or unique local problems, such as continuing in-migration of families with underimmunized children. In either case, shifting resources away from this county to others with “better” performance would be unlikely to result in improved immunization rates. At the same time, however, a more complete understanding of program performance and its relation to outcomes will support a more rational, albeit more complex, budgeting and resource allocation decision-making process.

The panel is concerned that some legislative actions to mandate performance standards and impose financial penalties for failure to comply make poor use of the performance measurement tool. For example, the 1992 Synar Amendment is intended to reduce tobacco consumption among youths by reducing their access to tobacco products. This provision requires that each state reduce to less than 20 percent the proportion of inspected sales outlets that violate the ban on the sale of tobacco products to those under age 18. States that repeatedly fail to meet the required level of performance face the loss of up to 40 percent of their Substance Abuse Block Grant funds (Substance Abuse and Mental Health Services Administration, 1998). Complicating the federal-state relationship on this issue are regulations issued by the Food and Drug Administration (1996) that make the sale of tobacco products to minors a violation of federal law, and preempt most state and local laws on this matter.

The panel sees at least four problems with the Synar Amendment’s approach to performance-based accountability. First, the performance requirement was established without states having the opportunity to participate as partners in identifying the performance measure to be used or the level of performance to be

achieved. Second, the financial penalty reduces the resources available to address prevention and treatment of all forms of substance abuse, not just youth tobacco use. Third, the performance requirement and its associated penalty are not related to the typical program goals and strategies of state substance abuse agencies. Few of these agencies have any enforcement authority regarding tobacco sales, and states are specifically prohibited from using their Substance Abuse Block Grant funds for any enforcement activities other than inspections of sales outlets. Finally, the penalty is based on a single process measure of performance (the proportion of sales outlets violating the ban on sales of tobacco to minors) without an assessment of the desired (intermediate) outcome—a reduction in tobacco use among minors—or conclusive evidence of a causal link between process and outcome (see Rigotti et al., 1997).

EXAMPLES OF PERFORMANCE MONITORING AND ACCOUNTABILITY SYSTEMS

The PPG proposal that served as the impetus for the work of this panel is but one application of the performance monitoring and accountability systems that are currently in use in a variety of settings. Perhaps the most prominent governmental example is GPRA, which requires all federal executive branch agencies to implement a strategic planning and performance measurement process. Various federal programs that provide funding to states also include performance reporting requirements. The Maternal and Child Health Bureau (MCHB) of the Health Resources and Services Administration in DHHS (1997) has incorporated performance measures into the reporting requirements for the agency's block grant. The new welfare block grant program, Temporary Assistance for Needy Families (TANF), links both penalties and bonus funds to state performance in specified areas. States will also be required to develop and report on performance measures in connection with the Children's Health Insurance Program, a major initiative to extend health insurance to currently uninsured children in low-income families who are not eligible for Medicaid.³ And many state governments are adopting performance-based management and budgeting systems. In the private sector, interest in assessing and improving the quality of health care is prompting the development of performance measurement systems for health plans, health care facilities, and individual health care providers. Some of these examples of the use of performance monitoring and accountability systems are reviewed briefly below.

³Information about the State Children's Health Insurance Program is available from the Health Care Financing Administration at <<http://www.hcfa.gov/init/children.htm>>.

Government Performance and Results Act

In 1993, Congress passed GPRA (P.L. 103-62) as part of an effort to improve the management and accountability of federal agencies. GPRA requires each agency to develop a strategic plan covering a period of at least 5 years, as well as annual performance plans and annual performance reports. Because GPRA requires major changes in agency management activities, its implementation is being phased in over several years. Agencies were required to submit their first strategic plan to the Office of Management and Budget (OMB) and Congress in September 1997. Annual performance plans were submitted beginning in 1998 for fiscal year 1999, and the first performance reports are to be issued in March 2000.

Each of the agency reporting requirements contributes to the overall performance-based management system envisioned under GPRA. The agencies' strategic plans are the starting point for defining program goals and outlining strategies for achieving those goals. Agencies are expected to consult with Congress and other stakeholders to ensure that their views are taken into consideration. The annual performance plan translates the broader, longer-term goals of the strategic plan into more operational goals for the coming year. Included in the annual performance plan are the performance measures the agency will use to assess progress toward its goals. In the annual performance report produced the following year, an agency is to use data collected for its performance measures to compare actual performance against the program goals. The aim over time is for these reports to include data for the reporting year plus the 3 prior years.

As agencies have been working with OMB and Congress to implement GPRA, GAO and a panel of the National Academy of Public Administration have produced several reports reviewing progress, noting problems, and recommending steps to support the implementation process (e.g., National Academy of Public Administration, 1994, 1998; U.S. General Accounting Office, 1996, 1997a,b). All agree that GPRA provides a sound framework and has the potential to bring substantial improvements to the management of federal programs. There are, however, significant challenges to be overcome if GPRA is to be successful.

In particular, GAO (1997a) has identified several problem areas that are hindering agency progress toward implementing the provisions of GPRA. The initial strategic planning step has proven difficult for some agencies because fragmented or overlapping programs are not easily translated into clear statements of agency mission and strategic goals. For some agencies, the challenge lies in reconciling competing or conflicting policy demands. GAO also suggests that there has been limited progress in the adoption of a results-oriented organizational culture to guide agency management decisions.

Of particular relevance to the work of this panel are GAO observations regarding the use of performance measures. Agencies are finding it difficult to measure performance on an annual basis when the outcome of program activities

cannot be determined within a single year, or the federal contribution to a result is only one of many influences on the outcome of interest. As the present panel did in its first report, GAO suggests compensating for these factors by using measures of intermediate results, using multiple measures, and working with stakeholders to agree upon measures to be used. Formal program evaluations can provide additional insight, but since they require substantial planning, time, and funding, it will be necessary to use them selectively (U.S. General Accounting Office, 1997b). GAO has also found that agencies lack suitable data for some performance measures because data for this purpose are not collected or are not of acceptable quality. For some agencies, a lack of baseline data has made it difficult to establish annual performance goals.

Use of Performance Measures in the Maternal and Child Health Services Block Grant Program

Within DHHS, MCHB has responded to the new GPRA requirements in part by introducing a performance measurement component into the Maternal and Child Health Services Title V Block Grant to States Program (see Maternal and Child Health Bureau, 1997, 1998a). Originally authorized in 1935 by Title V of the Social Security Act, this block grant provides federal funds to assist states in developing and operating programs intended to improve the health of pregnant women and children and provide services for children with special health care needs, including children with developmental disabilities or chronic illnesses. Four broad categories of services are supported: direct health care services (e.g., prenatal care), enabling services (e.g., case management, transportation), population-based services (e.g., immunizations, lead screening), and infrastructure-building services (e.g., needs assessments, information systems). States are required to match the block grant funds at a rate of \$3 for every \$4 in federal funds. Federal funding for fiscal year 1999 is \$580 million.

Since 1989, state accountability for the use of Title V funds has been linked to reporting on key maternal and child health indicators, as well as budget and expenditure data. Prompted by GPRA's new requirements for performance-based accountability and reporting by federal agencies, MCHB has revised the states' block grant reporting requirements to include the use of performance measures. In addition to providing a more effective indication of the impact of Title V programs at the state level, the state performance reports are expected to furnish data that MCHB will need to prepare GPRA performance reports for Congress.

Beginning with fiscal year 1998, each state must report on 6 health outcome measures (perinatal mortality, infant mortality, neonatal and postneonatal mortality, child death rates, and a measure of the disparity between black and white infant mortality rates), 18 "national" performance measures to be used by all states, and 7-10 additional performance measures selected by the state and approved through negotiation with MCHB. The negotiation process also includes

reviewing annual performance targets to be set by states for each measure. Examples of the 18 national measures are the percentage of a state's children with special health care needs who have a medical/health home,⁴ the birth rate for teenagers aged 15-17, the percentage of newborns screened for hearing impairment before hospital discharge, and the percentage of very low birth weight infants delivered at facilities for high-risk deliveries and neonates. Within this set of measures, all four categories of grant-supported services are represented.

Although the Title V Block Grant was not part of the original PPG proposal, MCHB has drawn on the PPG model in developing its performance measurement program. The Bureau has worked in partnership with the states, through the Association of Maternal and Child Health Programs and other channels, to reach agreement on the outcome measures and 18 national performance measures to be used. The outcome measures represent long-term health improvement goals to which Title V programs should be contributing but generally do not control. The national and state-selected performance measures are a mix of capacity, process, and risk factor measures and are linked more directly to program activities and shorter-term goals. This use of a mix of measures is consistent with the approach advocated in this panel's first report (National Research Council, 1997).

MCHB has also recognized that performance reporting should take into account differences among the states in their health needs and priorities and in the role Title V programs may play in meeting those needs. The use of state-selected measures allows states to emphasize program activities of special interest or importance. MCHB reviews these measures with each state to help ensure that they are practical and effectively link program activities and outcome goals. The review also gives MCHB an opportunity to increase the cross-state comparability of these data by encouraging states that select similar measures to adopt identical definitions of the numerators and denominators for those measures. States have access to technical assistance for their performance measurement work through MCHB offices and outside consultants. In addition, an MCHB systems development initiative is providing state grants of up to \$100,000 that can be used to support information systems activities related to Title V performance measurement (Maternal and Child Health Bureau, 1998b).

⁴Children with special health care needs are defined as "those who are at increased risk for chronic physical, developmental, behavioral, or emotional conditions, and who require health and related services of a type or amount beyond that required by children generally" (Maternal and Child Health Bureau, 1997:107). MCHB uses the definition of a medical/health home developed by the American Academy of Pediatrics (1992): a medical/health home is a source of accessible, continuous, comprehensive, family-centered, coordinated, and compassionate care provided by a well-trained physician able to manage or facilitate most forms of pediatric care.

Reporting Requirements for the Temporary Assistance for Needy Families Program

Although not a health program, the new federal program for public assistance to needy families is another example of the shift from categorical to block funding that gives states greater flexibility in return for accountability for their performance. The TANF program was enacted under the Personal Responsibility and Work Opportunities Reconciliation Act of 1996 (P.L. 104-193) (see Administration for Children and Families, 1998a). TANF replaces the Aid to Families with Dependent Children (AFDC) program and the Job Opportunities and Basic Skills (JOBS) training program. The new program aims to provide time-limited assistance to needy families and to reduce their dependence on government benefits by promoting job preparation, work, and familial responsibility through marriage. States have greater flexibility than under previous programs to specify who receives benefits, under what terms, and for how long, but they must submit reports demonstrating that their performance is in compliance with the provisions of the legislation and achieving desired outcomes.

The principal TANF performance standard relates to the work requirements for assisted families: a specified proportion of adult recipients must be engaged in work or allowable work-related activities for a minimum number of hours per week. For example, in 1998, 30 percent of all TANF families had to have an adult working at least 20 hours per week, with higher rates of work participation required of two-parent families. By 2002, 50 percent of all TANF families must have an adult working at least 30 hours per week. Evidence of substantial reductions in caseloads can substitute for achievement of the targeted work participation requirements. States must file quarterly reports to the federal government on these work participation rates.

TANF links both penalties and bonuses to the level of performance. States that do not meet the work participation requirements or other performance standards are subject to a reduction in their annual block grants. An initial penalty of 5 percent for noncompliance with work participation rates can be increased by 2 percent per year to a maximum of 21 percent for repeated noncompliance. States are, however, given the opportunity to develop a plan for achieving compliance before penalties are assessed. States also can compete for annual "high performance" bonuses intended to reward accomplishments in moving welfare recipients into jobs (Administration for Children and Families, 1998b). In the first year, states with the best performance on each of four measures of employment gains will be eligible for bonus awards.⁵ Because states can be rewarded based

⁵Four measures are to be used in selecting those states that will be eligible to receive high performance bonus awards: (1) the job entry rate for adult recipients of assistance; (2) success in the workforce, measured using the job retention rate and the earnings gain rate for employed adult recipients; (3) the annual percentage change in the job entry rate; and (4) the annual percentage change in the success-in-the-workforce measure (Administration for Children and Families, 1998b).

on the quality of the work that recipients find as well as the proportion of recipients who find work, the bonuses provide those states having less vibrant labor markets with an incentive for improvement. The TANF program also includes provisions for annual bonuses to the states that are most successful in reducing rates of out-of-wedlock childbearing.

On the basis of its observations regarding health-related block grant programs, this panel urges careful monitoring of the measures used to assess performance under TANF to ensure that they produce useful information without promoting unintended adverse effects. Some observers have expressed concern that these measures do not provide an adequate picture of program outcomes in terms of potential changes in the health and well-being of children in families receiving assistance or of adults or children in families that leave the welfare rolls (National Research Council, 1998).⁶ National and state data systems may need to be modified to produce such information. The panel also notes the limited opportunity states appear to have had to influence the performance criteria on which penalties are based. Greater collaboration is evident in determining how the high-performance bonus funds will be awarded. DHHS is working with the National Governors' Association, the American Public Human Services Association (formerly the American Public Welfare Association), and state representatives to develop the measures and formula to be used for this purpose (Administration for Children and Families, 1998b).

State Developments in Performance-Based Budgeting

States, like the federal government, are looking to revamp their program management process to better ensure desired outcomes for their citizens (Zelio, 1997). Current state-level performance monitoring and budgeting initiatives are the latest in a series of efforts to increase the responsiveness of state executive agencies to the electorate and the legislature. These initiatives seek to move beyond line-item budgeting, with its focus on detailed categories of expenses and resultant micromanagement of complex organizations, to an emphasis on program outcomes. Such efforts are generally driven by management-oriented state legislatures whose members believe that the implementation of improved management controls within state government systems will lead to more effective government overall.

A 1996 study found that 45 states are using performance measures in various ways (Florida Office of Program Policy Analysis and Government Accountability, 1997). There are 6 states using performance information as a budget decision tool, and another 9 are in the process of implementing such a system. Another

⁶A panel of the Committee on National Statistics of the National Research Council was convened in September 1998 to examine data and methods for measuring the effects of changes in social welfare programs. This panel's report should be completed in late 2000.

recent review of performance-based budgeting found that 7 states have introduced links between performance and financial or management incentives (e.g., financial rewards for agencies or individual employees, increased flexibility in use of funds); 2 of these states include disincentives such as increased oversight and reporting requirements (Melkers and Willoughby, 1998). States that engage in performance budgeting are actively restructuring their budget documents, reordering organizations, and changing organizational missions to align with policy responsibilities. In some instances, the organizations and suborganizations are realigned to be consistent with program objectives so that policy responsibilities are located within a single organization.

Oregon's strategic planning effort illustrates the use of a participatory state-wide approach to planning and setting performance goals (Oregon Progress Board, 1997). In a process that was initiated by the governor in 1989, a strategic plan for the state, not just state government, was developed with input from the public and private sectors, including the general public. A set of benchmarks⁷ was chosen to translate the goals of the strategic plan into measurable objectives on such matters as health, education, employment, and the environment. For example, current health-related benchmarks include measures such as the percentage of adults who do not currently smoke tobacco; the percentage of eighth grade students who used alcohol in the previous month; and the percentage of Oregonians with a lasting developmental, mental, and/or physical disability who work. These benchmarks have been used by the legislature and state agencies in setting program and budget priorities for which specific performance measures are developed. However, it may be difficult to relate a benchmark based on a summary measure (e.g., years of potential life lost before age 70) to specific program activities or funding needs.

A recent review of the state's strategic plan (Oregon Progress Board, 1997) resulted in several recommendations that may be relevant for other performance measurement activities. Among these recommendations was identifying the relationships among benchmarks (e.g., teen pregnancy and child poverty). In addition, the system should use benchmarks for which reliable data are regularly available at a reasonable cost. The targets selected for benchmarks should also reflect realistic, evidence-based expectations of achievable performance; for example, limitations in current understanding of the factors that affect birth weight make it unreasonable to set a target of reducing the number of low-weight births by 50 percent. Moreover, as responsibility for implementing programs is transferred to the community level, the development of accurate and timely local-level data

⁷Oregon uses the term "benchmark" to refer to the specific measures used to monitor progress toward the state's goals. Others use the term to refer to a selected standard against which performance is compared (National Performance Review, 1997). For example, a state attempting to reduce a high rate of adult smoking might adopt the lower national rate as a benchmark against which to assess progress. Oregon is using performance "targets" to assess progress on each measure.

becomes a priority. Another recommendation was to reduce the number of benchmarks from 259 to about 100, as the larger number of measures had proven difficult to track and prioritize. This panel notes, however, that a reduction in the number of measures involves a trade-off since it may lead to reduced visibility of some concerns within important specific areas (e.g., health).

Today's state-level performance monitoring and budgeting efforts vary in their focus, with their approach depending on which of three overall purposes they serve (Florida Office of Program Policy Analysis and Government Accountability, 1997):

1. *Guide management and administration.* This purpose is served by an approach, similar to that of GPRA, that relies on the stakeholders for each agency and its subagencies developing a long-range plan and defining outcome and other performance measures. These formalized measures are used to guide the management and administration of the organization. Although the measures may be shared externally with the legislative body or the public, their primary purpose is to help agencies focus on a particular set of goals.

2. *Inform the budget process.* States emphasizing this purpose concentrate their efforts on explaining the focus of their program and its achievements to the legislature and the public. The information provided is highly descriptive and includes details on capacity, resources, and expenditures. It allows legislators to make policy decisions in a larger context and consider the functions of all sectors of government. There are no direct financial or statutory incentives or disincentives under this approach.

3. *Provide a basis for resource allocation.* States with this focus hope that performance budgeting systems will provide the major rationale for allocation of funds and make it possible to set measurable objectives. An attempt is made to report on past performance and shift the focus from line budgets to desired outcomes. In exchange for accountability, these states hope to offer executive agencies flexibility in management as a way of rewarding achievement. While some reporting is reduced or eliminated, the approach incorporates periodic program-specific evaluations that are supported by independent scientific verification of performance to validate accountability.

State experiences with performance-based budgeting suggest several lessons (U.S. General Accounting Office, 1994; Florida Office of Program Policy Analysis and Government Accountability, 1997). States have found it important to involve a broad range of stakeholders in the strategic planning process to achieve consensus on program goals and measures. Legislative and executive leadership are needed to ensure continuity of objectives over time and continued availability of the resources necessary to produce performance measurement reports. There are major challenges involved in designing performance monitoring systems that can clearly define governmental responsibilities and are meaningful to decision

makers. Furthermore, despite their commitment to performance-based management, state government personnel generally need more training in the development and use of performance measures. Information systems are recognized as necessary components of a performance-based management system, but they are frequently inadequate to generate the needed data on outcomes, program processes, and strategy-specific costs.

Health Care Performance Measurement in the Private Sector

Until fairly recently, performance-based accountability for health care outcomes has operated primarily on a case-by-case basis through malpractice claims and quality assurance programs, reflecting an assessment of the care provided by individual clinicians or hospitals to individual patients. More recently, quality improvement and performance measurement programs have altered this accountability framework by introducing continuous monitoring of the processes and outcomes of care for populations of patients. As under the performance-based budgeting approaches described above, clinical performance information provides management tools that can be used to promote improvements in health care.

Some of the best-known recent efforts to develop performance measurement systems in health care have been led by employer groups, credentialing organizations, health maintenance organizations, hospitals, and private consultants. Among the leading private-sector efforts are those by the National Committee for Quality Assurance (NCQA), the Foundation for Accountability (FACCT), and the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), which are described in more detail below. In addition, the American Medical Association (1998) has introduced an accreditation program for individual physicians that will include standardized measures of clinical performance and patient care results.

There is an increasing degree of collaboration among these groups in the development of clinical performance measures and performance measurement systems. Moreover, as a growing proportion of Medicare and Medicaid services are provided by private-sector health plans, there is increasing public-private collaboration in the further development of some of these performance measurement systems. The federal government (U.S. Department of Health and Human Services, 1998) has announced plans to implement the recommendation of the President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry (1998) to establish a Forum for Health Care Quality Measurement and Reporting that will work with private-sector groups to develop a core set of measures and standards for measurement. The American Medical Accreditation Program of the American Medical Association, JCAHO, and NCQA have established a Performance Measurement Coordinating Council to coordinate their performance measurement activities and through which they anticipate working with the newly proposed forum (Joint Commission on

Accreditation of Healthcare Organizations, 1998b). In a more targeted collaboration, the Diabetes Quality Improvement Project has brought together FACCT and NCQA, plus the American Academy of Family Physicians, the American College of Physicians, the American Diabetes Association, the Health Care Financing Administration, and the Veterans Health Administration, to develop a set of diabetes-related performance measures suitable for use nationally (Diabetes Quality Improvement Project, 1998).

National Committee for Quality Assurance

One of the most prominent performance measurement tools in health care is NCQA's Health Plan Employer Data and Information Set (HEDIS), a set of standardized measures for comparing the quality of care provided by participating health maintenance organizations (National Committee for Quality Assurance, 1997a). Originally developed to inform employers purchasing health services for their employees, HEDIS has evolved to address consumer information needs as well. It now includes measures specifically for the Medicare and Medicaid populations, as well as the commercially insured. Health plan reports are filed with NCQA, which in 1997 began publishing an annual summary (National Committee for Quality Assurance, 1997e).

HEDIS 3.0, the most recent version, includes 71 measures that health plans are required to use and 32 other measures (a "testing set") that are undergoing further evaluation and refinement. Each measure has a standard definition and technical specifications for data collection and calculation. For the measures based on data to be obtained through a member satisfaction survey, a standardized survey instrument has been developed (National Committee for Quality Assurance, 1997b). The HEDIS 3.0 measures cover the following domains of performance: effectiveness of care, access/availability of care, satisfaction with the experience of care, health plan stability, use of services, cost of care, informed choice, and health plan descriptive information.

An ongoing review and development process has been established to support the continued evolution of HEDIS measures and the overall HEDIS system. The Committee on Performance Measurement, which oversaw the development of HEDIS 3.0, will continue to guide the review of current measures, the identification of measures to be retired, the testing of new measures, and a research agenda to support the development of new measures and overall improvements in performance measurement. Measurement advisory panels will provide additional expertise for work in specific areas (e.g., behavioral health, cardiovascular disease, women's health).

HEDIS has become a widely recognized set of performance measures for assessing health care services provided by health maintenance organizations, but some have found it too limited in certain areas. In particular, the limited number of measures on mental health and substance abuse services has led to efforts by

others to develop suitable measures for managed behavioral health services (e.g., American Managed Behavioral Healthcare Association, 1995; J. Dilonardo, Substance Abuse and Mental Health Services Administration, personal communication, 1998).

In an initial test of the feasibility of HEDIS, the Report Card Pilot Project provided useful lessons that were reflected in the development of HEDIS 3.0 (Spoeri and Ullman, 1997) and are relevant to the work of this panel. Specifically, the pilot project revealed the need to adopt a broad set of measurement domains and to field test measures before formal adoption. It also demonstrated the variation in the organization and operation of health plan information systems and the need for greater standardization to produce comparable data across plans. Clinical information systems were generally found to be weaker than those for administrative and financial data. External data audits were valuable in identifying errors and inconsistencies in data systems and in the specifications and processes used to calculate measures. The documentation for HEDIS 3.0 includes a set of audit standards (National Committee for Quality Assurance, 1997d) and a report specifically on the development of information systems that can support performance measurement using HEDIS (National Committee for Quality Assurance, 1997c). A continuing area of concern is the need for risk adjustment of HEDIS measures. Although this need has been recognized, suitable risk adjustment techniques for use across plans have not yet been developed.

Foundation for Accountability

FACCT was created in 1995 in response to a desire by consumer groups and purchasers of health care services for a more effective means of bringing their perspectives to bear on the assessment of health care quality (Foundation for Accountability, 1998a). Working with consumer focus groups and experts, FACCT has developed sets of measures for use in assessing care for adult asthma, alcohol misuse, breast cancer, diabetes, and major depressive disorder (Foundation for Accountability, 1998b). In terms of the panel's framework, these sets include measures of process, risk status, and outcomes, including measures of satisfaction with care for the specific condition. FACCT has also developed a set of measures that focuses on smoking as a health risk factor. Two other sets address general health status and overall consumer satisfaction with services and care (e.g., getting needed services, choice of providers). Under development are measurement sets for coronary artery disease, end-of-life care, HIV/AIDS, and pediatric care. The measures adopted by FACCT are field tested by health plans and group practices as part of the development process.

FACCT has placed special emphasis on the consumer perspective and seeks to measure elements of health care quality that are important to consumers. In recent work with the Health Care Financing Administration, FACCT (1997) developed a framework intended to communicate health care performance infor-

mation (e.g., measures from FACCT and HEDIS) to Medicare beneficiaries in an effective manner.⁸ The project also explored conceptual and technical issues involved in constructing summary performance scores for health plans or health care providers.

Joint Commission on Accreditation of Healthcare Organizations

JCAHO has long served as one of the principal accrediting bodies for health care facilities. Its accreditation programs now include hospitals, home care agencies, long-term care facilities, behavioral health services, ambulatory health care providers, laboratories, and health care networks. Efforts over the past few years to integrate clinical performance measurement into JCAHO's accreditation process resulted in the Oryx initiative, which began in 1997 (Joint Commission on Accreditation of Healthcare Organizations, 1998c). Included in the Oryx program are hospitals, long-term care organizations, health care networks and health plans, home care organizations, and behavioral health care organizations. In the past, the accreditation process has been based on evidence of compliance with JCAHO standards covering such matters as staff credentials, equipment, and policies (Joint Commission on Accreditation of Healthcare Organizations, 1998a). In the panel's performance measurement framework, these standards could be viewed as focusing primarily on capacity (i.e., inputs to health care services), rather than on processes or outcomes of care. The addition of performance measures is seen as a way for the accreditation process to stimulate and contribute to quality improvement efforts.

The Oryx program will allow health care organizations to meet their performance measurement requirements through the use of a variety of measurement systems. For hospitals and long-term care facilities, JCAHO has approved more than 200 measurement systems operated by a variety of organizations. These include JCAHO's own Indicator Measurement System, which offers a set of performance measures focused on specific areas of patient care (e.g., obstetrics, trauma, oncology). Measures for health care networks have been selected from measure sets developed by JCAHO, FACCT, NCQA, the University of Colorado Health Sciences Center, and the University of Wisconsin (Madison). Health care organizations will report their performance data through the organizations that manage the specific measurement systems they adopt, not directly to JCAHO.

To maintain their JCAHO accreditation, health care organizations must report on a specified minimum number of measures selected from approved measure-

⁸The following reporting categories are proposed: The Basics, covering elements such as access, provider skill, and communication; Staying Healthy, covering education, prevention, and risk reduction; Getting Better, covering treatment and follow-up for illness or injury; Living with Illness, covering functional status and quality of life for persons with chronic conditions; and Changing Needs, regarding end-of-life care or care at times of major changes in functional abilities.

ment systems. For example, hospitals and long-term care organizations must initially report on at least 2 clinical measures that together are relevant to at least 20 percent of their patient population, or they must report on 5 measures. Health care networks must initially report on 10 measures. Plans call for increasing the required number of measures and patient population coverage. Separate reporting requirements are being developed for each accreditation program.

An advisory council has been established to provide a continuing review of the measurement systems included in the Oryx program. This group will also help select a set of core measures for each accreditation program. Review of candidate measures for use by hospitals is expected to begin in late 1998. Recognizing that selection of a measurement system and use of specific performance measures will be unfamiliar tasks for some of the participating organizations, JCAHO has developed a guidebook and other resources to help organizations evaluate and select a measurement system that will meet their needs.

Lessons for Publicly Funded Health Programs

The evolution of performance measurement in health care in the private sector offers lessons to those developing performance measures for publicly funded health programs. One key lesson is that performance measurement requires a continuing effort to select and improve measures and the measurement process. The quality and usefulness of the performance data being produced by health care organizations continue to improve, but conceptual and technical challenges remain (see, e.g., Eddy, 1998). The individualized performance "report cards" developed in the past by some health plans lack the comparability across plans and providers that might be achieved by the larger-scale performance measurement programs, such as those of NCQA, FACCT, and JCAHO. These latter programs rely on more standardized sets of measures and guidelines for collecting relevant data using standard methodologies.

The activities of these nongovernmental groups are an important resource for performance measurement for the publicly funded health-related programs that the panel is addressing. The work done by these groups to identify suitable measures for clinical care can inform the selection of measures for related aspects of public programs. Likewise, the experience these groups are gaining in developing measurement standards and information system tools to support performance measurement in a health services context may help guide related efforts in the public sector.

CONCLUSIONS

Although the concept of performance measurement is hardly new and the use of performance indicators has been attempted episodically in various programs, the widespread use of such indicators in federal programs as contemplated by

GPRA is a new and significant requirement that is also emerging among state and local governments. Similarly, the increasingly widespread use of HEDIS and other performance measurement systems in health care is evidence of changing attitudes and expectations regarding accountability and management in the private sector.

Early experience with these vastly expanded requirements for accountability suggests that the new approaches offer many attractive features, but successful implementation will require substantial and continuing efforts to overcome several challenges. Conceiving and developing measures that capture performance accurately and comprehensively is often difficult and should be guided by special expertise; lack of data to support selected measures may necessitate the use of second-best choices; and multiple sets of measures may be required to satisfy the needs of varied users (e.g., program managers, funders, and the public). As more is learned about the use of performance measurement, progress is possible on all of these fronts. After reviewing performance measurement experience in other contexts, the panel concluded that several principles should guide current efforts to implement performance measurement for publicly funded health programs.

Link performance measurement to program goals. Performance measurement should be viewed as a tool that facilitates the monitoring and promotion of progress toward program goals, not as an end in itself. It must be based on a clear articulation of program goals and desired outcomes—health outcomes in the context of this report—and some sense of how those goals can be achieved. Outcome measures should reflect a program’s goals, and measures of process and capacity should reflect the evidence on effective methods of achieving those outcomes. Performance measurement should be a constructive process that contributes to organizational capacity to meet program goals.

Adopt a “market basket” approach. A performance measurement system should promote the development of recognized sets of measures with agreed-upon definitions from which program participants (e.g., states or communities) should be expected to select specific measures that reflect the program priorities and strategies they have adopted. Even though programs generally have a core set of goals and objectives that are applicable regardless of where the program is operating, they must respond to diverse needs and regional circumstances. This means that specific program priorities and the strategies adopted to achieve them are likely to vary across states and communities. Therefore, a single, mandated set of performance measures is not appropriate. However, an effort should be made to associate particular program goals and strategies with specific outcome, risk status, process, and capacity measures so that identical activities related to those goals and strategies can be monitored using the same measures. For example, a program to reduce teenage smoking might be expected to use a standard measure of smoking prevalence. The specific process and risk status measures adopted should reflect the choice of strategies for reducing the prevalence of teenage smoking (e.g., reducing access to cigarette vending machines, restrict-

ing tobacco advertising near schools). Ideally, each measure should be recognized as valid, reliable, and responsive to change.

Recognize differing needs for performance information. The content and number of useful performance measures should be expected to differ between a program's operating level and the policy and sponsorship level across the inter-governmental structure. Compared with other levels, the operating level is likely to require more measures, and measures that focus more on process than on outcome. A performance measurement system should recognize these differing needs, but aim to use measures that can be linked, conceptually or in practice, to provide a consistent assessment of performance across these different levels. This principle is consistent with the GAO (1996) recommendation regarding GPRA that "hierarchies" of performance goals and measures are needed to reflect differing roles and responsibilities at various organizational levels.

Ensure the feasibility of data collection and analysis. The most elegant performance measures are of little use without a feasible data system to support them. Considerations such as the quality of the available data and the cost of obtaining specific data elements may limit the choice of measures, particularly in the short run. In some cases, it may be necessary to use less desirable measures while enhancing existing data sources or building better data sets. The panel's first report (National Research Council, 1997) specifically noted that the lack of data comparable across states was a significant obstacle to identifying optimal performance measures for many program areas. Given the trade-offs involved, it is clearly important to consider data collection and analysis strategies as part of the development of performance measurement systems.

Assess the consequences of using performance measurement. Performance measurement may achieve the desired effect of improving outcomes, or it may inadvertently promote undesired effects. Measurement results could, for example, be misinterpreted. A state with rates of food-borne illness that are higher than those of other states could be viewed as having problems in food safety practices when, instead, the higher rates reflect a more effective surveillance system. Another undesirable effect might be neglect of program areas or activities that are not being measured. Prematurely high expectations for performance data or rapid adoption of rigid performance targets could undermine intended program goals. For example, program practices might be manipulated to achieve "good" results, perhaps by avoiding populations that are difficult to serve rather than by implementing more effective services. The performance monitoring system, including individual performance measures, should be evaluated periodically to assess the consequences of its use. Such evaluation would help ensure that the system's goals were being met and decrease the likelihood of manipulation or inadvertent adverse effects, such as reduced services to groups that may be likely to have poor outcomes.

Adopt a developmental approach. The development of a successful performance measurement system should be viewed as an activity that continues to

evolve over time. Furthermore, because performance measurement is a new and largely unfamiliar policy mechanism, it should be tested in the contexts of goal setting, progress monitoring, and signaling of progress or problems before being used for resource allocation or regulatory purposes. The panel advocates starting with a comprehensive vision for a performance measurement system that is implemented in manageable phases, during which the participants learn and the system grows. There must be a firm commitment to ongoing research to develop new and better measures, relate these measures to program actions, and improve the performance measurement system. Research and evaluation studies must be done to test the effectiveness of performance measurement as a tool for improving health outcomes and program management.

For health programs, measures should be refined or replaced as understanding of the linkages between health outcomes and program activities (“processes”) improves and as better sources of data are developed. Moreover, program priorities can be expected to change over time, necessitating the identification and testing of new performance measures. Sustained investments are needed in improvements to data systems, as well as in training and technical assistance to ensure that program and policy staff develop the necessary skills and expertise. With time and experience, performance measurement may prove to be an effective basis for allocation of resources or assessment of regulatory benchmarks, but it must always be used prudently, with an understanding of both its strengths and its limitations.

REFERENCES

Administration for Children and Families

- 1998a Temporary Assistance for Needy Families (TANF). Fact sheet. February 13, 1998. U.S. Department of Health and Human Services. <http://www.acf.dhhs.gov/programs/opa/facts/tanf.htm> (August 4, 1998).
- 1998b Formula for Awarding the First High Performance Bonus in Fiscal Year (FY) 1999. Memorandum to state agencies administering the Temporary Assistance for Needy Families (TANF) program and other interested parties. March 17, 1998. U.S. Department of Health and Human Services. <http://www.acf.dhhs.gov/news/welfare/highperf.htm> (August 4, 1998).

American Academy of Pediatrics

- 1992 The medical home. *Pediatrics* 90:774.

American Managed Behavioral Healthcare Association

- 1995 *Performance Measures for Managed Behavioral Healthcare Programs (PERMS)*. Washington, D.C.: American Managed Behavioral Healthcare Association.

American Medical Association

- 1998 American Medical Accreditation Program. <http://www.ama-assn.org/med-sci/amapsite/index.htm> (July 31, 1998).

Diabetes Quality Improvement Project

- 1998 Initial Measure Set (Final Version). August 14, 1998. <http://www.facct.org/DQIP-fnl.html> (November 25, 1998).

- Eddy, D.M.
 1998 Performance measurement: Problems and solutions. *Health Affairs* 17(4):7-25.
- Florida Office of Program Policy Analysis and Government Accountability
 1997 *Performance-Based Program Budgeting in Context: History and Comparison*. Report No. 96-77A (April). Tallahassee: Florida Legislature. Available at <http://www.opaga.state.fl.us/budget/reviews.html/> (April 8, 1998).
- Food and Drug Administration, U.S. Department of Health and Human Services
 1996 Regulations restricting the sale and distribution of cigarettes and smokeless tobacco to protect children and adolescents. Final rule. August 28. *Federal Register* 61(168): 44395-44445.
- Foundation for Accountability
 1997 Reporting Quality Information to Consumers. Portland, Ore.: Foundation for Accountability.
 1998a About FACCT. <http://www.facct.org/about.html> (April 15, 1998).
 1998b Measuring Quality. <http://www.facct.org/measures.html> (December 28, 1998).
- Institute of Medicine
 1997 *Improving Health in the Community: A Role for Performance Monitoring*. J.S. Durch, L.A. Bailey, and M.A. Stoto, eds. Committee on Using Performance Monitoring to Improve Community Health. Washington, D.C.: National Academy Press.
- Joint Commission on Accreditation of Healthcare Organizations
 1998a Accreditation Information. http://www.jcaho.org/acr_info/acr_std.htm (August 16, 1998).
 1998b Nation's Three Leading Health Care Quality Oversight Bodies to Coordinate Measurement Activities. Press release. May 19, 1998. <http://www.jcaho.org/news/nb.htm> (June 5, 1998).
 1998c Oryx Fact Sheet for Health Care Organizations. <http://www.jcaho.org/perfmeas/oryx/sidebar1.htm> (July 24, 1998).
- Maternal and Child Health Bureau
 1997 *Guidance and Forms for the Title V Application/Annual Report. Maternal and Child Health Services Title V Block Grant Program*. December 22, 1997. Rockville, Md.: U.S. Department of Health and Human Services, Health Resources and Services Administration.
 1998a Office of State and Community Health. U.S. Department of Health and Human Services, Health Resources and Services Administration. <http://www.hhs.gov:80/hrsa/mchb/osch.htm> (December 30, 1998).
 1998b State Systems Development Initiative (SSDI) Grant Application Guidance for FY98. U.S. Department of Health and Human Services, Health Resources and Services Administration. <http://www.hhs.gov:80/hrsa/mchb/guidance.htm> (June 4, 1998).
- Melkers, J., and K. Willoughby
 1998 The state of the states: Performance-based budgeting requirements in 47 out of 50. *Public Administration Review* 58(1):66-73.
- MHSIP Task Force on a Consumer-Oriented Mental Health Report Card
 1996 The MHSIP Consumer-Oriented Mental Health Report Card. Final report of the Mental Health Statistics Improvement Program (MHSIP) Task Force on a Consumer-Oriented Mental Health Report Card. Rockville, Md.: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- National Academy of Public Administration
 1994 *Toward Useful Performance Measurement: Lessons Learned from Initial Pilot Performance Plans Prepared Under the Government Performance and Results Act*. Washington, D.C.: National Academy of Public Administration.

1998 *Effective Implementation of the Government Performance and Results Act*. Washington, D.C.: National Academy of Public Administration.

National Committee for Quality Assurance

1997a *HEDIS 3.0/1998*. Washington, D.C.: National Committee for Quality Assurance.

1997b *HEDIS 3.0/1998*. Vol. 3, *Member Satisfaction Survey*. Washington, D.C.: National Committee for Quality Assurance.

1997c *HEDIS 3.0/1998*. Vol. 4, *A Road Map for Information Systems: Evolving Systems to Support Performance Measurement*. Washington, D.C.: National Committee for Quality Assurance.

1997d *HEDIS 3.0/1998*. Vol. 5, *HEDIS Compliance Audit Standards and Guidelines*. Washington, D.C.: National Committee for Quality Assurance.

1997e *The State of Managed Care Quality*. Washington, D.C.: National Committee for Quality Assurance. <http://www.ncqa.org/news/report.htm> (April 7, 1998).

National Performance Review

1993 *From Red Tape to Results: Creating a Government That Works Better and Costs Less*. Washington, D.C. Available at <http://www.npr.gov/library/nprprt/annrpt/redtpe93/index.html> (August 27, 1998).

1997 *Serving the American Public: Best Practices in Performance Measurement*. Washington, D.C. Available at <http://www.npr.gov/library/review.html> (July 23, 1998).

National Research Council

1997 *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health*. E.B. Perrin and J.J. Koshel, eds. Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics. Washington, D.C.: National Academy Press.

1998 *Providing National Statistics on Health and Social Welfare Programs in an Era of Change*. Summary of a workshop. C.F. Citro, C.F. Manski, and J. Pepper, eds. Committee on National Statistics. Washington, D.C.: National Academy Press.

Oregon Progress Board

1997 *Oregon Shines II: Updating Oregon's Strategic Plan*. Salem, Ore.: Oregon Progress Board.

Osborne, D., and T. Gaebler

1992 *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector*. Reading, Mass.: Addison-Wesley.

President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry

1998 *Quality First: Better Health Care for All Americans*. Washington, D.C.: President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. Available at <http://www.hcqualitycommission.gov/final/> (July 27, 1998).

Rigotti, N.A., J.R. DiFranza, Y.C. Chang, T. Tisdale, B. Kemp, and D.E. Singer

1997 The effects of enforcing tobacco-sales laws on adolescents' access to tobacco and smoking behavior. *New England Journal of Medicine* 337:1044-1051.

Spoeri, R.K., and R. Ullman

1997 Measuring and reporting managed care performance: Lessons learned and new initiatives. *Annals of Internal Medicine* 127:726-732.

Substance Abuse and Mental Health Services Administration

1998 SAMHSA's Tobacco Activities: Implementing the Synar Requirements Under the Substance Abuse Prevention and Treatment Block Grant. U.S. Department of Health and Human Services. <http://www.samhsa.gov/csap/synar/sydex.htm> (May 18, 1998).

U.S. Department of Health and Human Services

1991 *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. DHHS Pub. No. (PHS) 91-50212. Washington, D.C.: Office of the Assistant Secretary for Health.

1998 President Endorses Quality Commission's Final Report and Issues Executive Memorandum to Improve Health Care Quality. White House fact sheet. March 13, 1998. <http://www.hhs.gov/news/press/1998pres/980313a.html> (August 4, 1998).

U.S. General Accounting Office

1994 *Managing for Results: State Experiences Provide Insights for Federal Management Reforms*. GAO/GGD-95-22. Washington, D.C.: U.S. Government Printing Office.

1996 *Executive Guide: Effectively Implementing the Government Performance and Results Act*. GAO/GGD-96-118. Washington, D.C.: U.S. Government Printing Office.

1997a *The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven*. GAO/GGD-97-109. Washington, D.C.: U.S. Government Printing Office.

1997b *Managing for Results: Analytic Challenges in Measuring Performance*. GAO/HEHS/GGD-97-138. Washington, D.C.: U.S. Government Printing Office.

1997c *Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation*. GAO/AIMD-97-46. Washington, D.C.: U.S. Government Printing Office.

Wholey, J.S.

1983 *Evaluation and Effective Public Management*. Boston: Little, Brown.

Wholey, J.S., and H.P. Hatry

1992 The case for performance monitoring. *Public Administration Review* 52(6):604-610.

Zelio, J.

1997 Update on performance budgeting. *LegisBrief* 5(37). National Conference of State Legislatures.

3

Performance Measurement Considerations for Publicly Funded Health Programs

As discussed in the preceding chapter, performance measurement is a prominent feature of current policy and management approaches. This panel's efforts have been focused on performance measurement in the broad array of health-related programs supported in some measure by public funding. In its first report, the panel addressed primarily the federal-state funding relationship for the specific set of programs included in the performance partnership grant (PPG) proposal (see Chapter 1). In this report, the discussion has been expanded so that many of the concerns addressed are relevant for performance measurement more generally, not just in the context of a federal-state funding relationship. The previous chapter explored how performance measurement is being used in some of these other settings (e.g., for federal agencies reporting to Congress, for state agencies reporting to their legislatures, and for health plans in the private sector).

The discussion here has also been broadened to look beyond the specific program areas covered by the panel's first report. Each program area poses unique performance measurement challenges. These challenges reflect factors such as the nature of the services and program activities being undertaken, the extent to which evidence of effectiveness is available to guide program activities, and the degree of consensus on appropriate measures to be used. For example, the control of food-borne disease outbreaks generally requires much more rapid response than do steps to control cancer, and such differences should be reflected in the performance measures used.

Although program-specific issues must be considered, the panel emphasizes that a strictly programmatic perspective may discourage a more comprehensive approach to performance measurement that can capitalize on the complementary,

overlapping, and even synergistic interactions among programs and their information system needs. Thus, the panel has attempted to consider a mix of specific and cross-cutting issues. This chapter reviews several of these issues, including the broad array of health-related services and service relationships, measurement considerations for population-based health services, special considerations in specific health program areas, and the importance of using process guidelines as a basis for performance measurement.

BROAD ARRAY OF HEALTH-RELATED SERVICES AND SERVICE RELATIONSHIPS

The panel's work has focused on performance measurement for health-related programs for which public funding is provided and for which performance-based accountability is sought to foster effective use of resources. Yet the programs for which this accountability is sought cannot be viewed only in terms of a particular funding arrangement. The stated goals and objectives of a health program, not the source of funding, should be the primary focus of performance evaluation. Both the breadth and the limitations of the services and service relationships involved are important because those elements must be accommodated in the performance measurement systems that are developed. Four key considerations are noted here.

First, most health program areas that receive public funding are influenced by a diverse array of factors. Funds may come from federal, state, local, and private sources and from different service categories within those sources. As the panel emphasized in its first report, this means that program outcomes must generally be viewed as the collective product of all these contributions and can rarely be credited to a single funding source. Thus, even though a single funder, such as a federal block grant program, may establish a requirement for performance measurement, the development of measures and related data resources requires consideration of the full scope of influences on the program being assessed.

Second, the focus on publicly funded program areas means the panel concentrated on performance measurement and data system issues of concern to public agencies at the federal, state, and local levels. Relevant work in the private sector, such as the health care performance measurement activities discussed in Chapter 2, should be taken into consideration, but the panel has not attempted to formulate recommendations regarding those activities.

Third, implementation of publicly funded health programs involves not only public-sector health agencies, but also agencies with other responsibilities (e.g., education, criminal justice, housing, transportation) and organizations and individuals in the private sector, such as hospitals, health plans, individual clinicians, and employers. This means that the planning and implementation of performance measurement for health programs should take a broad view of the stakeholders involved.

Finally, the health-related programs the panel has considered provide a variety of services, ranging from clinical care for illness to population-based services aimed at health promotion and disease prevention. Mental health programs, for example, typically emphasize clinical care to treat individual patients, whereas chronic disease programs are more likely to offer screening services, such as cholesterol testing for an entire community. A variety of programs may make use of public education aimed at the community at large. Environmental services, such as water treatment and restaurant inspections, are essential for protecting the health of all members of the community, but are not delivered directly to individuals. The mix of such services varies widely among programs. Using expenditures as a measure, the Public Health Foundation (Eilbert et al., 1996) found that mental health agencies devote almost all of their resources to personal health care services, whereas environmental agencies, which have major nonhealth responsibilities, support primarily population-based health services. Performance measurement efforts should have links to all these kinds of services and consider their differing data collection needs and data resources.

MEASUREMENT CONSIDERATIONS FOR POPULATION-BASED HEALTH SERVICES

In developing performance measurement systems for the broad range of publicly funded health programs, certain factors will be relevant to most, if not all, programs. For example, all programs should be monitored using a mix of capacity, process, and outcome measures, and those measures should be as valid, reliable, and responsive to the changes they are expected to monitor as possible. Further, these measurement systems must respond to changing health needs, measurement tools, and program resources. Within a state or community, several programs may benefit by coordinating both their services and their information systems.

Population Health Services

Many publicly funded health programs fall within the realm of “public health,” a designation based not on the source of funding or on the specific content of the services but on the population-based approach used to plan and provide those services. The defining features of public health are its emphasis on protecting and improving the health of the general public through prevention-oriented population-based services, and its role in ensuring that key services reach individuals at risk. Examples include programs to provide clean water, immunizations for children, and adequate prenatal care to the disadvantaged. Because the health of the public, or of specific groups in the population, is influenced by a mix of factors, some of which are beyond the control of the individuals involved, public health programs often depend on collective action by various institutions

of society to achieve the full potential of health in a community. For purposes of performance measurement, public health activities with population-wide goals and objectives should be recognized as having goals distinct from those for personal health services, which focus on care for individuals.

Population health services are based on a public health perspective that focuses on an assessment of the overall health needs of the population. Some services (e.g., water treatment, public education programs, tobacco control) are provided to the population at large; others are delivered directly to individuals (e.g., immunizations, family planning, screening services) as a way to improve both individual health and the overall health status of the population. In contrast, personal health services are based on a clinical perspective that focuses on the care sought by an individual (e.g., diagnosis of disease, a surgical procedure, counseling services for participants in a substance abuse treatment program). Collectively, these personal health services for individuals contribute to better health for the population as a whole, but personal health services are not specifically intended to meet population health goals.

Population health services have important interrelationships with personal health services. Both may play a role in providing services for primary prevention or responding to certain health problems, and the benefits may be realized by specific individuals and the population in which they live. For example, the timely diagnosis and successful treatment of a case of infectious tuberculosis cures the individual and prevents the infection from spreading in the population. Moreover, diagnosis of an individual case of tuberculosis can trigger a systematic screening of population groups that may be at increased risk for infection. Other examples of this synergy can be found in the benefits for both individuals and society of successful treatment of those who abuse alcohol and other drugs, which can help reduce problems such as vehicle-related injuries caused by drunk driving, domestic violence, and crime related to illegal drugs. Such interconnections between personal and population health services are reflected in the recently renewed appreciation of the value of collaboration between the domains of medical care and public health (Lasker et al., 1997).

The distinctions between population and personal health services have implications for performance measurement and monitoring. For population health services, health outcomes and risk status are measured by overall changes for a population (or subgroup) as a whole. For personal health services, interventions must be monitored on the basis of the response of those individuals who received the services.

Monitoring Population Health Services

Many public health agencies at the local, state, and federal levels have an established foundation of ongoing collection of health-related data (e.g., vital records, infectious disease reporting, cancer registries, surveys on health status

and risk behaviors, hospital discharge reports) to inform programs, policy makers, and the public about the health status of the population and the effectiveness of health programs. In general, these data systems are oriented to producing information about the health of the population rather than to tracking the health of specific individuals. These activities, often referred to collectively as public health surveillance, are a key component of public health services. Public health surveillance has been defined as

the ongoing systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who have a need to know. The final link in the surveillance chain is the application of these data to prevention and control. A surveillance system includes a functional capacity for data collection, analysis, and dissemination linked to public health programs (Thacker and Berkelman, 1988:164).

Plans for performance measurement in publicly funded health programs should build on the surveillance systems already in place. They are a primary source of data for performance measures used to assess programmatic activities, especially measures of health outcomes and risk status. Other data collection activities may be needed to produce data for measures of program processes and capacity.

Those population health services that are generally regarded as highly successful in the prevention of adverse health outcomes may present special challenges for performance measurement. The protective effects of such services are often taken for granted, but their failure has the potential for widespread and serious consequences in the population. For example, a major outbreak of cryptosporidiosis in Milwaukee in 1993 occurred when essential water treatment systems broke down. For these services, outcome measures such as the incidence of disease are most informative only when some aspect of the system fails, not when it is functioning properly. Therefore performance measurement, in addition to monitoring health outcomes such as disease incidence, should focus on the steps taken to protect against such failures. These protective steps are generally best represented by capacity and process measures, such as water chlorination levels and numbers of inspections, that provide indications of appropriate risk reduction practices.

Some population health services exert an indirect influence on health or contribute to positive outcomes for future generations. They may act through the collective action of community groups (for examples, see Institute of Medicine, 1996a), and they may require sustained efforts to achieve the desired outcomes. The performance measures to be used to monitor such services must be selected carefully. As discussed earlier, the short time horizon usually adopted for performance monitoring—a period of from 3 to 5 years was proposed for PPGs—may dictate the use of intermediate outcomes because the longer-term outcomes can-

not be observed within the specified time frame. Process and capacity measures might be designed to assess the collaboration and continuity needed to achieve those longer-term outcomes.

An additional concern is ensuring that performance measurement promotes, or does not hinder, the “equity” of population health services. The use of measures that focus only on the total population can obscure problems among high-risk populations, which might be defined by geography, race or ethnicity, or risk-related characteristics. Program goals and the associated performance measures should be framed in a way that gives attention to all relevant populations, and data collection, particularly through surveys, must be designed to produce statistically meaningful performance measurement results for those population groups.

Monitoring the Infrastructure for Publicly Funded Health Programs

The Future of Public Health (Institute of Medicine, 1988) describes the core functions of public health agencies at all levels of government as *assessment* of community (or population-wide) health status and health needs, *policy development* to protect and promote the health of the public, and *assurance* that services necessary to achieve health goals are provided. In recent years, public health practitioners have identified a set of 10 “essential services” that describe how the three core functions of public health are carried out (see Box 3-1).¹

These core functions and essential services might be considered part of the infrastructure that supports all publicly funded health programs. Performance measurement itself is readily encompassed as a responsibility of health agencies through both the assessment and assurance functions and at least two of the essential services: monitor health status to identify community health problems, and evaluate effectiveness, accessibility, and quality of personal and population-based health services.

The panel encourages states and communities to consider using performance measurement to monitor not only the programmatic aspects of public health services—immunization programs, water treatment, and maternal and child health programs, for instance—but also the infrastructure for these health programs as represented by activities related to the core functions and essential services.

For example, monitoring the accuracy and completeness of surveillance services is important because insufficient or poorly conducted surveillance may

¹The list of essential public health services was adopted by the Public Health Functions Steering Committee in 1994 as part of its statement *Public Health in America*. This group coordinates the Public Health Functions Project, which was created to clarify issues related to the public health infrastructure and develop strategies and tools to address those concerns. The steering committee is chaired by the Assistant Secretary for Health and the Surgeon General, and its members include Public Health Service agency heads and presidents of several national organizations representing state and local health officials and other public health practitioners.

Box 3-1
Essential Public Health Services

1. Monitor health status to identify community health problems.
2. Diagnose and investigate health problems and health hazards in the community.
3. Inform, educate, and empower people about health issues.
4. Mobilize community partnerships to identify and solve health problems.
5. Develop policies and plans that support individual and community health efforts.
6. Enforce laws and regulations that protect health and ensure safety.
7. Link people to needed personal health services and assure the provision of health care when otherwise unavailable.
8. Assure a competent public health and personal health care workforce.
9. Evaluate effectiveness, accessibility, and quality of personal and population-based health services.
10. Research for new insights and innovative solutions to health problems.

SOURCE: Public Health Functions Steering Committee, 1994.

miss health problems or provide a misleading picture of health status, especially in comparison with results based on good surveillance data. In the short term, a community or population that is served by inadequate surveillance may inaccurately be perceived as “healthier” than a community with a more comprehensive surveillance system that detects more cases of illness. This effect was seen when a state that quickly and accurately collected, analyzed, and interpreted data on salmonella infections detected a major outbreak, producing the impression that the population was less healthy than those of other states (Van Beneden et al., 1996). Further investigation showed, however, that similar outbreaks had gone undetected in other states with less proficient surveillance systems, making their populations appear healthier than they were.

Although the panel supports implementing performance measurement to monitor public health services, additional groundwork will be needed to reach agreement on an approach to measurement and to identify suitable measures and data sources. Furthermore, research and evaluation remain necessary to determine the impact on health outcomes of the performance of activities related to the core functions and essential services of public health.

Until a clearer consensus emerges regarding such measures, states and communities may wish to refer to some of the tools that have been developed to help

local health departments assess their ability to perform the core public health functions and deliver associated services (e.g. National Association of County Health Officials, 1991; National Civic League, 1993; Centers for Disease Control and Prevention, 1995). Other work to develop formal measures of effective local health department performance may also be informative (Miller et al., 1994a,b; Turnock et al., 1994a,b, 1995; Richards et al., 1995). Proposals for objectives for public health infrastructure for Healthy People 2010 (Office of Disease Prevention and Health Promotion, 1997) may also suggest performance measures for this purpose. Additional guidance for such performance measures should be expected from efforts to establish a national accreditation program for local health departments and to develop national public health performance standards (Halverson et al., 1998). This work is being done collaboratively by the Centers for Disease Control and Prevention (CDC), the Association of State and Territorial Health Officials, the National Association of County and City Health Officials, the National Association of Local Boards of Health, the Public Health Foundation, and the American Public Health Association.

SOME PERFORMANCE MEASUREMENT CONSIDERATIONS RELATED TO PROGRAM-SPECIFIC MATTERS

Many health program areas have distinctive features that will need to be taken into account as performance measurement systems are developed. Some of these features for selected program areas are reviewed in this section.

Environmental Health Programs

Environmental health services, such as ensuring clean air, safe water, and protection from other toxic exposures, are a classic component of public health programs, but they pose difficult challenges for performance measurement. Environmental threats to health and the services designed to control those threats are diverse. For example, environmental health services must address both the low-level air and water contamination that can exacerbate conditions such as asthma and increase the longer-term risks for cancers, chronic respiratory illness, and other adverse health outcomes. They must also address the sudden high-level toxic exposures that produce acute health effects requiring immediate medical attention. The primary goal is to prevent both types of exposure and to ensure that if either should occur, an effective means of responding will be available. Environmental health outcomes of specific exposures reflect interactions among personal susceptibility, other hazards in the environment (synergistic effects), the biologically effective doses of the hazards, and the mitigating effects of protective measures that may operate by affecting any of these elements (National Research Council, 1994).

Efforts to monitor environmental health risks and steps taken to control them

therefore require a mix of information on the hazards (e.g., specific air or water pollutants), the exposures (e.g., biological markers, such as blood lead levels), and health outcomes (e.g., asthma, birth defects, and cancer) (see, e.g., Thacker et al., 1996). Such monitoring activities can be challenging because of the variety of potential hazards, the range of settings in which exposures can occur (e.g., home, workplace, community at large), the differing sources of hazardous exposure (e.g., multiple small sources, such as vehicle emissions or pesticide runoff in surface water, versus point sources, such as smokestack emissions or wastewater discharged from a manufacturing plant), and the interaction between private-sector business interests and public-sector regulation. Moreover, if health effects occur long after an exposure or when an exposure is not perceived, health care providers treating individual patients may not be able to identify the link between the current health problem and the past exposure. Growing interest in environmental equity and environmental justice have also drawn attention to the need to address the disproportionate exposure to environmental hazards (e.g., hazardous waste sites, hazardous manufacturing processes) faced by certain communities, neighborhoods, and other population groups.

The public health community is working to develop better ways of addressing these environmental health challenges. For example, the National Association of County and City Health Officials (1997), with support from the National Center for Environmental Health at CDC, is working to develop methods for use by communities in assessing environmental health conditions. Environmental health data issues have also been the focus of various workshops (e.g., National Center for Environmental Health, 1996; Public Health Foundation, 1997).

This work has highlighted several problem areas that are of particular concern for performance measurement efforts. Scientific evidence on the links between environmental exposures and health outcomes is limited. Even where a risk exists and these links are understood, available data may not be adequate to assess the exposures of individuals or population groups. Data are collected by a variety of federal, state, and local agencies and in the private sector by individual companies, but lack of coordination in these data collection activities can lead to redundancies and to inconsistencies across data systems. It can also be difficult to identify and gain access to the data, and proprietary data may never come to public attention. A crucial factor for performance measurement is a lack of consensus on appropriate indicators of environmental health status or of capacity and processes in environmental health services.

The Environmental Protection Agency (EPA) is using performance measurement in its performance partnerships with states and has undertaken several initiatives aimed at strengthening environmental data more generally (see Environmental Protection Agency, 1998a,b). EPA notes that these collaborative efforts with states and industry include a reassessment and reorganization of reporting requirements to reduce unnecessary reporting and achieve better coordination, the development of data standards, and the modernization of informa-

tion systems. Because of the diversity of interests served by environmental data, the panel emphasizes the need for federal, state, and local health agencies to ensure that they are represented in these EPA discussions so that health data requirements and concerns receive appropriate attention.

Mental Health Programs

Until the mid-1960s, most publicly funded mental health services were provided in institutional settings. Frequently, individuals with serious and persistent mental illness (e.g., schizophrenia and bipolar disorder) were treated in state institutions for many years—sometimes for their lifetimes. The Community Mental Health Center Act of 1963 promoted the development of community-based programs as an alternative to institutional care. New treatment strategies and the relatively recent development of more sophisticated medications have also made it possible for increasing numbers of people who were previously hospitalized to be cared for in community-based settings.

With a greater proportion of publicly funded mental health services being delivered in nonhospital settings and the recent trend toward managed care contracting, the nature of the service delivery system is changing. Increasingly, these publicly funded services are being provided by either for-profit or not-for-profit contractors. Such changes in the past few years have increased the importance of using performance contracts and outcome measures to respond to a demand for higher levels of accountability and the importance of having surveillance systems in place to monitor the impact of those changes. There is also an increased demand that such measures address the outcomes of interest to consumers and their families.

Until recently, there has been little consensus on how performance should be assessed. States, provider organizations, and accrediting bodies have undertaken separate efforts to develop their own standards and measures. Among the groups developing performance evaluation systems are the American College of Mental Health Administrators, the American Managed Behavioral Healthcare Association, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the National Committee for Quality Assurance (NCQA), the National Alliance for the Mentally Ill (NAMI), and the National Association of State Mental Health Program Directors (NASMHPD). At the federal level, the Center for Mental Health Services of the Substance Abuse and Mental Health Services Administration (SAMHSA) and the Health Care Financing Administration (HCFA) are also actively involved. The health outcome and risk status measures proposed in the panel's first report are listed in Appendix A.

All of these efforts have helped advance performance evaluation, but they have not resulted in the standard measurement system that was envisioned for the PPG proposal. Currently, almost every state mental health agency uses some set of measures to evaluate the impact of mental health services. Most states have

developed their performance evaluation packages by means of a local consensus-building process that has included consumers, advocates, and providers. The result typically is a system that is customized for the state, but with little similarity to the performance evaluation systems used in other states.

Increasingly, the emphasis is on the development of a common outcome-oriented framework for the evaluation of mental health programs (e.g., MHSIP Taskforce on a Consumer-Oriented Mental Health Report Card, 1996; Institute of Medicine, 1997b; Smith et al., 1997; National Association of State Mental Health Program Directors, 1998). The Mental Health Statistics Improvement Program (MHSIP) Consumer-Oriented Mental Health Report Card (MHSIP Taskforce on a Consumer-Oriented Mental Health Report Card, 1996) was one of the initial efforts to develop a measurement framework that incorporates outcomes of mental health services and specifically considers consumer concerns. An evolving consensus is reflected in the December 1997 adoption by NASMHPD members of a standardized performance indicator framework for the evaluation of public mental health services (National Association of State Mental Health Program Directors, 1998). The National Association of State Mental Health Program Directors Research Institute (1998b) is including indicators from the NASMHPD framework in the performance measurement system it is developing for use by state psychiatric hospitals to fulfill their accreditation requirements under the JCAHO (1998) Oryx program (see Chapter 2 for discussion of the Oryx program).

The NASMHPD framework draws from other efforts, such as the MHSIP Report Card and the work of the American College of Mental Health Administrators, NAMI, and NCQA. The framework was also influenced by a collaborative review of measures used in managed care settings conducted by SAMHSA; HCFA; and state mental health, substance abuse, and Medicaid programs. The five performance domains of the NASMHPD framework (and examples of the associated indicators) include access (enrollment rate, utilization rate), quality/appropriateness (consumer participation in treatment planning, contact within 7 days following hospital discharge), outcome (school improvement, symptom relief), structure/plan management (consumer/family member involvement in policy development), and early intervention/prevention (substance abuse screening, use of self-help/self-management).² Formal specifications for performance measures to operationalize these indicators are being prepared. The domains of the MHSIP Report Card are similar, but exclude structure/plan management.

The panel believes that further advances in performance-based accountability for public mental health systems will depend on four factors. First, consensus must be reached among stakeholders (e.g., consumers; clinicians; health plans;

²The domains of access, quality/appropriateness, and outcome can be considered equivalent to the domains identified in the panel's first report as important in the evaluation of mental health services. Those domains are, respectively, access to and utilization of services, quality assurance, and consumer satisfaction with services and psychological and social outcomes.

program managers; and federal, state, and local mental health agencies) on the domains that should be addressed by the performance measurement process. The adoption of the NASMHPD framework by the state mental health directors is an important advance.

Second, agreement must be reached on the concerns to be addressed in each domain and on the specific measures to be used to represent those concerns. The specification of outcome measures, including ones that reflect the consumer's perspective, is considered a particularly high priority. Examples of the indicators proposed in the NASMHPD (1998) framework to reflect outcome-related concerns are symptom relief, employment or school status, consumer perception of outcomes, and living situation.

Work is also needed on process and capacity measures for assessing access to care and the appropriateness of care. Such measures have been widely used, but have generally not been selected on the basis of evidence that they are linked to desired outcomes—a key requirement that the panel emphasized in its first report. This disconnect between outcomes and other measures points to the panel's third concern: research and program evaluation studies are needed to build a stronger evidence base linking mental health outcomes to specific aspects of process and capacity.

The fourth area that will require attention if performance measures are to be used successfully is the further development of agreed-upon data collection tools and procedures and their integration into existing mental health program information systems. A recent study to test the ability of five states to use many of the measures identified in the MHSIP Report Card and the NASMHPD framework—the Five-State Feasibility Study—found that fewer than half of the 28 measures tested could be reported by all five states and that differing definitions frequently limit the comparability of apparently similar measures (National Association of State Mental Health Program Directors Research Institute, 1998a). A subsequent study will build on this work with a group of 10 states. The performance measurement requirement for states' JCAHO-accredited hospitals is also likely to help stimulate further development of these measurement tools and information systems.

In contrast to the emphasis on population-based preventive services in many public health programs, mental health programs are focused almost entirely on services for persons with mental disorders. In some areas, however, a population perspective is useful. Assessments of the overall prevalence and incidence of mental disorders in the general population could help mental health programs gauge the potential need for individual or community-based services. Evaluation of preventive interventions and interventions outside the treatment setting (e.g., reducing the number of people with serious mental illness in jails and prisons) will also require the use of population-based measures. Suitable measures and data collection instruments will have to be developed and tested. The Epidemiologic Catchment Area Study (Bourdon et al., 1992) and the National Comorbidity

Survey (Kessler et al., 1994) have provided data for national prevalence estimates, but these studies were not designed to produce data on an ongoing basis or for states and communities, which performance measurement will require. CDC is working with several researchers to explore whether questions on the annual, state-based Behavioral Risk Factor Surveys can be used to obtain valid assessments of the mental health status of the population (Centers for Disease Control and Prevention, 1998).

In summary, the growing demand for performance-based evaluations of publicly funded mental health programs is creating an urgent need to develop greater consensus on the overall framework for these evaluations. This framework should define the domains for assessment (e.g., outcomes, appropriateness of and access to services), indicators that identify the critical concerns in each domain, and an array of specific measures and measurement tools (e.g., specific instruments for client assessment). As in other program areas, users should have the flexibility to select instruments and measures that best match program goals and strategies.

Substance Abuse Programs

Prevention and treatment of substance abuse are a high priority at the national, state, and local levels. Substance abuse includes use of illegal drugs, as well as inappropriate use of legal products such as alcohol and prescription medications.³ Substance abuse demands attention from a health perspective not only because it produces serious and difficult-to-treat physical and psychological effects, but also because it substantially increases the risk of other health problems, such as injury, adverse pregnancy outcomes, tuberculosis, HIV infection, and sexually transmitted diseases. In contrast to many other health problems, substance abuse is also an important criminal justice issue because use of many abused substances is illegal and because substance abuse tends to generate other criminal activity that adversely affects the general population.

Substance abuse programs have a clear stake both in population-based activities such as health education aimed at prevention and in personal health services needed for treatment of substance abuse and the other health problems it generates. Treatment may be supplemented by wraparound services that help people function more effectively in the community (e.g., transportation, housing, job placement).

The implication for performance measurement is that substance abuse programs might be expected to address a variety of outcomes, ranging from the

³The panel's work on substance abuse in the first phase of the study and the discussion in this section of the current report focus on program activities related to drug and alcohol abuse. In some contexts, smoking and other forms of tobacco use may be viewed as substance abuse issues. The panel's first report addressed tobacco use in the context of chronic disease prevention. See Appendix A for the tobacco-related risk status measures proposed in that report.

impact of treatment on the social functioning of treated individuals to the prevalence of substance abuse in specific population groups (e.g., adolescents) that are often the focus of prevention efforts. As noted in the panel's first report (National Research Council, 1997), however, few potential substance abuse performance indicators are measured in exactly the same way by all states or other jurisdictions. The health outcome and risk status measures proposed for substance abuse programs in the panel's first report are listed in Appendix A. There is general agreement on the content areas of greatest interest (see Box 3-2), but there is substantial variation in the program strategies adopted and the characteristics of the populations served in specific settings (e.g., public versus private, managed care versus fee-for-service).

Several activities are under way that, over time, are expected to contribute to greater consistency in the measures used. In October 1997, a workgroup organized by the National Association of State Alcohol and Drug Abuse Directors (NASADAD) proposed the adoption of a performance measurement framework based on the domains of efficiency, effectiveness, and structure (see National Association of State Alcohol and Drug Abuse Directors, 1998). Indicator areas and possible measures or data sources have been proposed for each domain. The proposed indicators for effectiveness are physical and mental health status, economic self-sufficiency, social supports and functioning, and alcohol and other drug use. The proposed indicators for efficiency are access, treatment retention, costs of services, and appropriateness of services. For structure, the proposed indicators are service capacity, data capabilities, workforce competence, and

Box 3-2
Performance Measurement Domains for Substance Abuse
Identified in Phase I PPG Process

- Treatment effectiveness
- Treatment completion
- Medical screening
- Use during pregnancy
- HIV- and STD-related
- Overall use and consequences
- Youth use and consequences
- Other prevention activities
- Access and special needs
- General and infrastructure issues

SOURCE: National Research Council (1997).

client characteristics. NASADAD is coordinating a discussion among its full membership to refine this proposal and achieve consensus on the various components. As the process moves forward, detailed specifications will be developed for individual measures. NASADAD is working closely with SAMHSA in this activity and is consulting with NASMHPD in areas of common interest.

SAMHSA's Center for Substance Abuse Treatment (CSAT) is involved in various activities related to performance measurement. For example, the Treatment Outcomes and Performance Pilot Studies project has funded 14 states to test methods of monitoring the performance of publicly funded substance abuse treatment services. Another priority is the development of measures that can be used in managed care settings. Currently, the Health Plan Employer Data and Information Set (HEDIS) (National Committee for Quality Assurance, 1997) is one of the primary assessment tools for managed care services, but it includes few measures related to substance abuse treatment services. In March 1998, CSAT and other SAMHSA units with responsibilities in managed care began discussions with a small group of providers, researchers, federal and state policy makers, and representatives from public- and private-sector managed care organizations aimed at identifying additional measures that might be used. Plans call for further discussions with a broader group of participants to refine the proposed measures and promote consensus. Other efforts include the development by the Foundation for Accountability (1998) of measures to assess health plan services to detect and treat alcohol misuse.

CSAT is also working on improving the availability of data to support performance measurement. Three states are testing the feasibility of integrating data related to substance abuse treatment from separate data systems operated by the state Medicaid, mental health, and substance abuse agencies. The project aims to produce a flexible model for this process that other states can apply to their specific organizational and data system configurations. In other work, CSAT is helping states develop data sets and information systems for monitoring treatment outcomes (e.g., Harrison, 1995). Much of this work, however, is in the context of evaluation studies that are essential to establish an evidence base for effective treatment services, but are not necessarily designed to produce data on a routine basis for performance monitoring.

The Center for Substance Abuse Prevention (CSAP) within SAMHSA has been working with several states to test a minimum data set on prevention services. The states reached consensus on five indicators that could be used for performance measurement: youth use (age at first or early use, current use), youth attitudes toward use, parental attitudes toward youth use, actual or perceived availability of specific substances, and ability to comply with Synar Amendment provisions on controlling the sale of tobacco to youth (the Synar Amendment is discussed in Chapter 2). NASADAD's discussions on performance indicators noted the need for prevention indicators, but NASADAD is

deferring work in this area until greater progress has been made on the treatment indicators.

A broader national performance measurement activity related to substance abuse is being led by the Office of National Drug Control Policy (ONDCP) (1998). Two of the five ONDCP goals—to prevent drug use among America’s youth and to reduce the health and social costs of drug use—specifically address health-related concerns. For each goal, several specific objectives have been established, and performance targets and associated measures have been chosen for both the goals and the objectives. For example, the goal of reducing the health and social costs of illegal drug use has six objectives: improve the drug treatment system, reduce drug-related health problems, promote a drug-free workplace, support training of the workforce for substance abuse services, develop medications and guidelines for substance abuse treatment, and support research and analysis to reduce the health and social costs of substance abuse. Among the performance measures adopted for this goal are the prevalence of drug abuse and the number of chronic drug users. Examples of the measures for the objective on improving the drug treatment system are the rate of full-time employment among adults completing substance abuse treatment programs and the average waiting time to enter treatment. Although ONDCP is focusing on national results, its activities may be sufficiently influential that states and communities will adopt measures that match those being used by ONDCP, thus achieving a greater consistency in measurement practices.

PROCESS GUIDELINES AS A BASIS FOR PERFORMANCE MEASUREMENT

As noted earlier, the panel concluded in its first report that performance monitoring requires the use of outcome measures and related process and capacity measures. The panel recommended that each process and capacity measure be accompanied by reference to published “guidelines or other professional standards that describe the relationship between the process or capacity measure and the desired health outcome” (National Research Council, 1997:2). The panel recognized, however, that such guidelines are not always available. In such cases, the panel recommended specifying the assumed relationship between proposed process or capacity measures and a health outcome, and documenting the assumed relationship with empirical evidence and professional judgment. Where guidelines are lacking, additional research is needed to establish more precisely the relationship between program interventions and outcomes. The panel recommended that DHHS sponsor empirical outcome studies so that a more definitive list of recommended process and outcome measures can be developed.

Similar recommendations regarding the use of evidence-based performance measures emerged from the Institute of Medicine (1997a) report *Improving Health in the Community*. This report addresses the use of performance monitoring in

community health improvement activities. It advises giving priority to health improvement actions that can be linked to evidence of effectiveness, but cautions that such evidence is limited for many health issues. It may be appropriate for communities to address those issues, but they must consider carefully which actions will make the best use of their resources.

Thus both reports point out the need for evidence concerning processes that lead to better health outcomes. This evidence is needed to guide performance as well as to design better performance measures.

Guidelines for Personal Health Services

Generally speaking, the evidence linking processes and outcomes is more extensive and more fully documented for personal health services than for population-based services. For instance, the *Guide to Clinical Preventive Services* (U.S. Preventive Services Task Force, 1996), first published in 1988, provides a rigorous assessment of evidence concerning the effectiveness of personal health services for disease prevention, such as screening, immunization, chemoprophylaxis, and health counseling, that are provided to individuals in clinical care settings. Between 1989 and 1996, the Agency for Health Care Policy and Research (AHCPR) sponsored the development of a series of 17 clinical practice guidelines on topics such as the use of mammograms for breast cancer screening, diagnosis and treatment of depression in primary care, and smoking cessation (Agency for Health Care Policy and Research, 1998). Clinical practice guidelines for many areas of clinical care have also been developed by a variety of groups, such as medical specialty organizations, insurers, and health care organizations. Reports from the Institute of Medicine (1990, 1992) provide a framework for promoting the development and use of high-quality guidelines.

AHCPR was also directed to develop clinical performance measures related to clinical practice guidelines. An Institute of Medicine (1990) report elucidated the distinction between guidelines and performance measures and the connections between the two. AHCPR (1995) subsequently published a working group report that describes how to construct performance measures related to individual guideline recommendations by identifying a population of individuals to whom the recommendation applies and then collecting data that show whether these individuals received the recommended care. Performance measures of this type, and specifically those related to the U.S. Preventive Services Task Force guidelines for clinical preventive services, appear in sets of performance measures for managed care organizations such as HEDIS (National Committee for Quality Assurance, 1997).

AHCPR's activities have shifted from the development of practice guidelines to support for Evidence-Based Practice Centers to develop reports on the scientific basis for interventions to prevent, diagnose, treat, and manage common diseases and clinical conditions (Agency for Health Care Policy and Research,

1998). These reports are intended to assist both public and private organizations in developing and implementing their own guidelines and performance measures. Among the topics chosen for the first round of evidence-based reports, some relate to areas considered by the panel, such as pharmacotherapy for alcohol dependence and treatment of depression with new drugs.

Guidelines for Population-Based Health Services

Practice guidelines and evidence-based reports, such as those developed by AHCPR, the U.S. Preventive Services Task Force, and other groups, often cover such topics as immunization and screening services that lie in the overlap between clinical care and public health. The development of guidelines and evidence-based reports for the population-based services that are at the heart of public health programs is just beginning. One of the major challenges in this process will be the limited availability of evidence regarding the effectiveness of community-based interventions. Studies of these interventions are difficult to design and conduct (see below for further discussion of this point).

Perhaps the most significant effort in this area is being undertaken by the Task Force on Community Preventive Services, which is working on a *Guide to Community Preventive Services* (U.S. Public Health Service, 1998). This guide (expected to be published in 2000) will complement the *Guide to Clinical Preventive Services* by focusing on community-based prevention and disease control strategies. It will provide evidence-based recommendations for interventions and their implementation. Separate sections of the guide will cover changing risk behaviors (e.g., tobacco use, sexual behavior, physical activity); reducing specific diseases, injuries, and impairments (e.g., vaccine-preventable diseases, violent behavior); and addressing environmental and ecosystem challenges. A section will also be devoted to cross-cutting public health activities such as surveillance.

Research Needs for Practice Guidelines and Performance Measurement

For practice guidelines and performance measurement in both clinical care and public health services, evidence is needed not only on *whether* an intervention works—whether it is causally associated with desired outcomes—but also on *how* it works. Moreover, the use of performance measurement must be studied to assess its effect on health outcomes and program operations.

Evidence on how a successful intervention works can be used to guide the organization, operation, and improvement of the associated services, and the selection and use of meaningful process and capacity performance measures. For example, evidence shows that prevention and treatment can be effective in reducing substance abuse, but further studies are needed to clarify which elements of these interventions contribute in what degree to successful outcomes (see Institute of Medicine, 1996b; McLellan et al., 1996, 1997; Landry 1997). For popula-

tion-based interventions, studies must also distinguish between the factors that contribute to success at the individual level and those that lead to successful outcomes for the population as a whole. For example, evaluations of several community-based programs designed to reduce coronary heart disease have shown that the programs were often successful in reducing disease risk for many individuals, but generally were not able to reach a sufficiently large proportion of the population to alter community-level health outcomes (e.g., Elder et al., 1993; Fortmann et al., 1995; Murray, 1995; Luepker et al., 1996). In addition, the unanticipated strength of other influences that were acting to reduce risks for heart disease largely overwhelmed the community-level impact of the interventions being tested.

Gaps in the available evidence can help indicate areas in which further research is needed. Such research will require a variety of approaches, including qualitative analysis (e.g., ethnographic studies) and quantitative analysis using techniques such as randomized controlled trials, quasi-experimental outcomes research, epidemiological studies, and program evaluation studies. The community-based interventions that are part of many publicly funded health programs have proven particularly challenging to study (see, e.g., Koepsell et al., 1992; Connell et al., 1995). For example, because communities are constantly affected by many factors other than the intervention of interest, it is difficult to identify an appropriate comparison (i.e., a control group or counterfactual) for judging what would have happened without the intervention. Of necessity, comparisons are often based on the community itself before the intervention or on one or two other communities considered similar in key characteristics. However, unanticipated changes in social or economic conditions or unidentified differences among communities can limit the usefulness of these comparison cases. The small number of communities included in most studies further complicates the analysis by making it difficult to detect a statistically significant community-level effect from the intervention or to conclude that the intervention has had no effect.

A major research effort will be necessary to establish a firm scientific basis for practice guidelines for individual and community-based interventions and for appropriate uses of performance measurement to monitor the implementation of those guidelines. The research in even a single area, such as smoking cessation, must cover practices as diverse as television advertising to inform the public about the nature of the risk and about aids to smoking cessation, enforcement of regulations against selling tobacco products to underage individuals, and bans on tobacco use in public places. The scale of research will vary from testing small programs in a few schools and workplaces to comparisons involving whole cities or states. In order to be effective, these research studies must have access to data on other factors in the community environment (e.g., changes in health care guidelines, changes in the local economy, natural disasters) that may affect the outcome of program efforts.

Studies must also be done to assess the effects of using performance measurement and to ensure that the commitment of resources to this activity is appropriate. It should not be assumed that performance measurement will have the desired positive effect on health outcomes and program management. Ideally, evaluations would demonstrate that performance monitoring activities contribute to protecting the health of the population by promoting such practices as the identification of and intervention against important health problems, coordination of information resources across health-related agencies, and the use of program practices consistent with evidence-based guidelines. Also valuable, however, would be learning that certain performance monitoring activities were inappropriately focused on health outcomes of minor significance to a population or subpopulation, leaving major health problems undetected or unaddressed.

As the fruits of this research effort become available, it will become possible to design more effective health programs, as well as to design better systems for monitoring performance within these programs.

CONCLUSIONS

Publicly funded health programs cover a broad spectrum of programmatic areas and include a diverse mix of population-based and personal health services. A performance monitoring system provides a framework both for defining desired health outcomes in each program area and for focusing attention on the steps being taken by health programs to achieve those outcomes. The panel cautions that given the complexity of influences on health outcomes, those outcomes can rarely be credited to a single program or funding source. Although the panel has focused on performance measurement for public-sector accountability, performance measurement can help ensure accountability to either public or private investors in and purchasers of these services.

Consideration of program-specific issues is important in defining performance measures and data needs. Consensus must be established regarding the appropriate domains of measurement and the measures to be used. Nevertheless, a strictly programmatic perspective could discourage a more comprehensive approach to performance measurement that can capitalize on the interrelationships among programs and the overlapping aspects of their data needs. For example, performance measurement systems might be developed for elements of the public health infrastructure such as surveillance systems. These infrastructure components contribute to essential public health services that can support efforts across a range of categorical program areas. Performance measurement in certain program areas (e.g., environmental health, mental health, substance abuse) requires involvement with and an understanding of programs outside of the publicly funded health arena, such as air quality management and criminal justice systems.

The panel emphasizes that performance measurement should rest on a strong

evidence base if links are to be established between desired health outcomes and both program activities (represented by process and capacity measures) and intermediate outcomes (represented by changes in risk status). Looking only at data on health outcomes may provide little insight into the contributions to good results being made by health program activities or into changes in those activities that might be necessary to improve health outcomes. Performance measurement draws attention to and establishes accountability for processes and intermediate outcomes that are more clearly under the control of health programs. Performance measurement may also promote the needed development of and adherence to evidence-based best practices to guide steps to achieve desired health outcomes, including a much-needed emphasis on defining standards of practice in public health program areas. In many program areas, the evidence base for performance measurement must be strengthened through additional research, including research to evaluate the effectiveness of performance monitoring itself.

REFERENCES

Agency for Health Care Policy and Research

1995 *Using Clinical Practice Guidelines to Evaluate Quality of Care. Volume 2: Methods.* AHCPR Pub. No. 950046. Rockville, Md.: U.S. Department of Health and Human Services, Public Health Service.

1998 Clinical Information. U.S. Department of Health and Human Services. <http://www.ahcpr.gov/clinic> (May 4, 1998).

Bourdon, K.H., D.S. Rae, B.Z. Locke, W.E. Narrow, and D.A. Regier

1992 Estimating the prevalence of mental disorders in U.S. adults from the Epidemiologic Catchment Area Survey. *Public Health Reports* 107(6):663-668.

Centers for Disease Control and Prevention

1995 *Planned Approach to Community Health: Guide for the Local Coordinator.* Atlanta, Ga.: National Center for Chronic Disease Prevention and Health Promotion.

1998 Self-reported frequent mental distress among adults—United States, 1993-1996. *MMWR* 47:325-331.

Connell, J.P., A.C. Kubisch, L.B. Schorr, and C.H. Weiss, eds.

1995 *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts.* Washington, D.C.: Aspen Institute.

Eilbert, K.W., M. Barry, R. Bialek, and M. Garufi

1996 *Measuring Expenditures for Essential Public Health Services.* Washington, D.C.: Public Health Foundation.

Elder, J.P., T.L. Schmid, P. Dower, and S. Hedlund

1993 Community heart health programs: Components, rationale, and strategies. *Journal of Public Health Policy* 14:463-479.

Environmental Protection Agency

1998a Administrator's Announcement on Reinventing Environmental Information. February 4, 1998. <http://www.epa.gov/reinvent/onestop/cbmemo6.htm> (April 28, 1998).

1998b One Stop Program Strategy and Grant Award Criteria. February 10, 1998. <http://www.epa.gov/reinvent/onestop/strategy.htm> (April 28, 1998).

Fortmann, S.P., J.A. Flora, M.A. Winkleby, C. Schooler, C.B. Taylor, and J.W. Farquhar

1995 Community intervention trials: Reflections on the Stanford Five-City Project experience. *American Journal of Epidemiology* 142:576-586.

Foundation for Accountability

- 1998 Measuring Quality. <http://www.facct.org/measures.html> (December 28, 1998).

Halverson, P.K., R.M. Nicola, and E.L. Baker

- 1998 Performance measurement and accreditation of public health organizations: A call to action. *Journal of Public Health Management and Practice* 4(4):5-7.

Harrison, P.

- 1995 *Developing State Outcome Monitoring Systems for Alcohol and Other Drug Abuse Treatment*. DHHS Pub No. (SMA) 95-3031. Rockville, Md.: Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment.

Institute of Medicine

- 1988 *The Future of Public Health*. Committee for the Study of the Future of Public Health. Washington, D.C.: National Academy Press.
- 1990 *Clinical Practice Guidelines: Directions for a New Program*. M.J. Field and K.N. Lohr, eds. Committee to Advise the Public Health Service on Clinical Practice Guidelines. Washington, D.C.: National Academy Press.
- 1992 *Guidelines for Clinical Practice: From Development to Use*. M.J. Field and K.N. Lohr, eds. Committee on Clinical Practice Guidelines. Washington, D.C.: National Academy Press.
- 1996a *Healthy Communities: New Partnerships for the Future of Public Health*. M.A. Stoto, C. Abel, and A. Dievler, eds. Committee on Public Health. Washington, D.C.: National Academy Press.
- 1996b *Pathways of Addiction: Opportunities in Drug Abuse Research*. Committee on Opportunities in Drug Abuse Research. Washington, D.C.: National Academy Press.
- 1997a *Improving Health in the Community: A Role for Performance Monitoring*. J.S. Durch, L.A. Bailey, and M.A. Stoto, eds. Committee on Using Performance Monitoring to Improve Community Health. Washington, D.C.: National Academy Press.
- 1997b *Managing Managed Care: Quality Improvement in Behavioral Health*. M. Edmunds, R. Frank, M. Hogan, D. McCarty, R. Robinson-Beale, and C. Weisner, eds. Committee on Quality Assurance and Accreditation Guidelines for Managed Behavioral Health Care. Washington, D.C.: National Academy Press.

Joint Commission on Accreditation of Healthcare Organizations (JCAHO)

- 1998 Oryx Fact Sheet for Health Care Organizations. <http://www.jcaho.org/perfmeas/oryx/sidebar1.htm> (July 24, 1998).

Kessler, R.C., K.A. McGonagle, S. Zhao, C.B. Nelson, M. Hughes, S. Eshleman, H.U. Wittchen, and K.S. Kendler

- 1994 Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry* 51(1):8-19.

Koepsell, T.D., E.H. Wagner, A.C. Cheadle, D.L. Patrick, D.C. Martin, P.H. Diehr, and E.B. Perrin

- 1992 Selected methodological issues in evaluating community-based health promotion and disease prevention programs. *Annual Review of Public Health* 13:31-57.

Landry, M.J.

- 1997 *Overview of Addiction Treatment Effectiveness*. Pub. No. DHHS(SMA) 97-3133. Rockville, Md.: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration.

Lasker, R.D., and the Committee on Medicine and Public Health

- 1997 *Medicine and Public Health: The Power of Collaboration*. New York: The New York Academy of Medicine.

Luepker, R.V., L. Rastam, P.J. Hannan, D.M. Murray, C. Gray, W.L. Baker, R. Crow, D.R. Jacobs, P.L. Pirie, S.R. Mascioli, M.B. Mittlemark, and H. Blackburn

- 1996 Community education for cardiovascular disease prevention: Morbidity and mortality results from the Minnesota Heart Health Program. *American Journal of Epidemiology* 144:351-362.

- McLellan, A.T., G.E. Woody, D. Metzger, J. McKay, J. Durell, A.I. Alterman, and C.P. O'Brien
1996 Evaluating the effectiveness of addiction treatments: Reasonable expectations, appropriate comparisons. *Milbank Quarterly* 74(1):51-85.
- McLellan, A.T., M. Belding, J.R. McKay, D. Zanis, and A.I. Alterman
1997 Can the outcomes research literature inform the search for quality indicators in substance abuse treatment? Pp. 271-311 in M. Edmunds, R. Frank, M. Hogan, D. McCarty, R. Robinson-Beale, and C. Weisner, eds, *Managing Managed Care: Quality Improvement in Behavioral Health*. Institute of Medicine. Washington, D.C.: National Academy Press.
- MHSIP Task Force on a Consumer-Oriented Mental Health Report Card
1996 The MHSIP Consumer-Oriented Mental Health Report Card. Final report of the Mental Health Statistics Improvement Program (MHSIP) Task Force on a Consumer-Oriented Mental Health Report Card. Rockville, Md.: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- Miller, C.A., K.S. Moore, T.B. Richards, and C. McKaig
1994a A screening survey to assess local public health performance. *Public Health Reports* 109:659-664.
- Miller, C.A., K.S. Moore, T.B. Richards, and J.D. Monk
1994b A proposed method for assessing the performance of local public health functions and practices. *American Journal of Public Health* 84:1743-1749.
- Murray, D.M.
1995 Design and analysis of community trials: Lessons from the Minnesota Heart Health Program. *American Journal of Epidemiology* 142:569-575.
- National Association of County Health Officials
1991 *APEXPH: Assessment Protocol for Excellence in Public Health*. Washington, D.C.: National Association of County Health Officials.
- National Association of County and City Health Officials
1997 Community Environmental Health Assessment: Project Fact Sheet. Washington, D.C.
- National Association of State Alcohol and Substance Abuse Directors
1998 Performance Measures. <http://www.nasadad.org/permeas.htm> (June 15, 1998).
- National Association of State Mental Health Program Directors (NASMHPD)
1998 State Mental Health Directors Adopt a Framework of Performance Indicators for Mental Health Systems. Press announcement. May 1998. Alexandria, Va.
- National Association of State Mental Health Program Directors Research Institute
1998a *Five State Feasibility Study on State Mental Health Agency Performance Measures*. Final Report. Alexandria, Va.: National Association of State Mental Health Program Directors Research Institute.
- 1998b NRI Performance Measurement System Home Page. <http://dmhmrs.chr.state.ky.us/nripms/> (August 14, 1998).
- National Center for Environmental Health
1996 Environmental Public Health Surveillance Workshops. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. <http://www.cdc.gov/nceh/programs/ephs/wkshop/summary.htm> (March 12, 1998).
- National Civic League
1993 *The Healthy Communities Handbook*. Denver: National Civic League.
- National Committee for Quality Assurance
1997 *HEDIS 3.0/1998*. Washington, D.C.: National Committee for Quality Assurance.
- National Research Council
1994 *Science and Judgment in Risk Assessment*. Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Studies and Toxicology. Washington, D.C.: National Academy Press.

- 1997 *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health.* E.B. Perrin and J.J. Koshel, eds. Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics. Washington, D.C.: National Academy Press.
- Office of Disease Prevention and Health Promotion
- 1997 *Developing Objectives for Healthy People 2010.* Washington, D.C.: U.S. Department of Health and Human Services.
- Office of National Drug Control Policy
- 1998 *Performance Measures of Effectiveness: A System for Assessing the Performance of the National Drug Control Strategy.* Washington, D.C.: Office of National Drug Control Policy.
- Public Health Foundation
- 1997 *Environmental Health Data Needs: An Action Plan for Federal Public Health Agencies.* Washington, D.C.: Public Health Foundation.
- Public Health Functions Steering Committee
- 1994 Public Health in America. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. <http://web.health.gov/phfunctions/public.htm> (August 31, 1998).
- Richards, T.B., J.J. Rogers, G.M. Christenson, C.A. Miller, M.S. Taylor, and A.D. Cooper
- 1995 Evaluating local public health performance at a community level on a statewide basis. *Journal of Public Health Management and Practice* 1(4):70-83.
- Smith, G.R., R.W. Manderscheid, L.M. Flynn, and D.M. Steinwachs
- 1997 Principles for assessment of patient outcomes in mental health care. *Psychiatric Services* 48:1033-1036.
- Thacker, S.B., and R.L. Berkelman
- 1988 Public health surveillance in the United States. *Epidemiologic Reviews* 10:164-190.
- Thacker, S.B., D.F. Stroup, R.G. Parrish, and H.A. Anderson
- 1996 Surveillance in environmental public health: Issues, systems, and sources. *American Journal of Public Health* 86:633-638.
- Turnock, B.J., A. Handler, W.W. Dyal, G. Christenson, E.H. Vaughn, L. Rowitz, J.W. Munson, T. Balderson, and T.B. Richards
- 1994a Implementing and assessing organizational practices in local health departments. *Public Health Reports* 109:478-484.
- Turnock, B.J., A. Handler, W. Hall, S. Potsic, R. Nalluri, and E.H. Vaughn
- 1994b Local health department effectiveness in addressing the core functions of public health. *Public Health Reports* 109:653-658.
- Turnock, B.J., A. Handler, W. Hall, D.P. Lenihan, and E. Vaughn
- 1995 Capacity building influences on Illinois local health departments. *Journal of Public Health Management and Practice* 1(3):50-58.
- U.S. Preventive Services Task Force
- 1996 *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force*, 2nd ed. Baltimore, Md.: Williams & Wilkins.
- U.S. Public Health Service
- 1998 Task Force on Community Preventive Services. The Guide to Community Preventive Services. <http://web.health.gov/communityguide> (January 1998).
- Van Beneden, C.A., W.E. Keene, D.H. Werker, et al.
- 1996 A Health Food Fights Back: An International Outbreak of Salmonella Newport Infections Due to Alfalfa Sprouts. Abstract K46, 36th ICAAC (Interscience Conference on Antimicrobial Agents and Chemotherapy). New Orleans, La. September 18, 1996.

4

Data and Information Systems: Issues for Performance Measurement

A performance measurement program must begin by identifying outcome goals, and then using those goals to guide the selection of suitable measures of desired outcomes and related processes and capacities. Once those steps have been completed, operationalizing performance measurement requires access to appropriate data and analytic resources. In its first report, the panel observed that many types of data useful for monitoring the performance of publicly funded health programs are collected and assembled across the country, but that few data sources are ideal for this purpose. For the most part, data systems have not been created specifically for performance measurement, so they may currently be narrower, less timely, or less comparable to other data systems than is optimal.

Despite such shortcomings, there are a number of reasons why the panel favors enhancing this extensive and often strong information base rather than establishing wholly new and specialized data systems for performance measurement. Although current health data collection processes and the resulting data sets often are not well coordinated with each other (Thacker and Stroup, 1994), the panel is hopeful that the current interest in performance measurement, reflected in reports such as this one, will encourage policy makers and health professionals at the federal, state, and local levels to transform the many different existing data sources into a more efficient and effective health information system with the capability of responding to varied information needs.

Collecting and assembling data is expensive, and expanding data collection efforts carries the risk of reducing the resources available for program services. Building on existing data systems for purposes of performance measurement would still require a substantial commitment of resources, but should be expected

to promote more efficient and effective use of those systems, and to improve their value for other applications as well. In relying on data collected for other primary purposes, however, those who develop and use performance measures must have a good understanding of the nature and limitations of those data.

This chapter begins by reviewing various health data resources. It then examines analytic and operational challenges involved in using those data, including assuring the quality of data and data analysis; developing and implementing standards for data and data systems; enhancing performance measurement through advances in information technology; and protecting the privacy, confidentiality, and security of health data. The chapter then outlines steps that can be taken to strengthen the data and data systems used to support performance measurement, in particular by investing in health data and data systems and by taking a collaborative approach to their development.

HEALTH DATA RESOURCES

Diverse health-related data are required to monitor and better understand the health of the population, including the incidence and prevalence of disease, morbidity and mortality associated with acute and chronic illness, behavioral risk factors, disability, and the quality of life. Data are also needed to plan, implement, and evaluate health policies, programs, and services. The data to meet these needs are produced and used in both the public and private sectors and, increasingly, by public-private partnerships. The Performance Partnership Grants (PPG) proposal that was the impetus for the work of this panel focused attention specifically on data for performance measures to be used in the context of state reporting requirements for federal grants. The panel emphasizes, however, that if performance measurement activities are to succeed, they should fit into a broader agenda for collecting and using health data to protect the health of the public, as well as for guiding the development and implementation of health policies at the local, state, and federal levels.

Although the panel did not attempt to address measurement of the quality and performance of individual health care providers or health plans, it should be noted that these activities are generating similar concerns about such matters as the selection of suitable performance measures, the limitations of administrative data sets for assessing health outcomes, the need for greater standardization of measures and data and for methods to improve data quality, and broader use of new information technologies (see, e.g., Iezzoni, 1997a; National Committee for Quality Assurance, 1997; Palmer, 1997; Foundation for Accountability, 1998; and Joint Commission on Accreditation of Healthcare Organizations, 1998). Major changes in social welfare programs are also prompting a reexamination of the adequacy of data resources for monitoring those programs, especially at the state and local levels (e.g., Joint Center for Poverty Research, 1998; National Research Council, 1998).

Data for performance measurement can be drawn from a variety of sources, such as reports to disease surveillance or vital statistics systems, environmental monitoring systems, population surveys, and clinical or administrative records from service encounters. Considering only the program areas covered by the original PPG proposal, the panel identified 48 data systems that might provide data for performance measurement (National Research Council, 1997a). Most states and communities can be expected to have a similarly large number of systems from which to draw data for performance measurement.

Four basic types of data resources are available: (1) *registries*, often referred to as census data systems, that attempt to capture information about all events of interest on such matters as health status (e.g., births, deaths, cases of disease) or risk factors (e.g., immunizations, environmental contaminants); (2) *surveys* that obtain data through the systematic collection of information from a representative sample of a population of interest; (3) *patient records* that contain clinical information obtained in the course of providing health care; and (4) *administrative data*, such as billing records, that are collected as part of the operation of a program (although these records may include data on health status or clinical care, that is not their primary purpose). Each type of data has a place in performance measurement, but each also has limitations that must be taken into account. Linking data over time or across data sets can potentially overcome some of those limitations and result in more useful information than is obtainable using a single data set or data for a single point in time. The basic types of health data and some of the issues related to linkage of data sets are reviewed in this section.

Registries

Registries are census-like data systems designed to compile information on all events of a specified type, such as births, deaths, specific injuries and environmental or infectious diseases, cancers, immunizations, hospital discharges, and birth defects. Vital records and disease surveillance registries are some of the most long-standing examples of these health data systems. Reporting systems also compile information on air and water quality, work-related injuries, and motor vehicle crashes resulting in deaths. Registries rely on reports of specific information to a designated authority. Some registries collect data through direct reporting of the events of interest (e.g., births, cases of reportable diseases), whereas others rely on assembling information originally collected in whole or in part for other purposes (e.g., work-related injuries).

Some of these systems operate locally, while others are connected to a state- or nationwide data system. For example, hospitals file reports on births with local or state registrars, and states then transmit these records to the National Center for Health Statistics (NCHS), where national vital statistics data are compiled. The rules governing which data are collected and how they are reported are developed and maintained through a federal-state collaborative system. In

contrast, immunization registries are being developed by some states and communities to capture reports on all immunizations administered to children (and also to serve as an information resource for health care providers on the immunization status of children under their care), but there is no national registry of immunization reports.

Registry systems benefit from standardized reporting practices. For example, NCHS and the states work together to develop standard birth and death certificates and guidelines for completing them. Systems differ substantially in their completeness, however. For example, virtually all births are reported, but reporting of fetal deaths is much less complete. Data on some reportable but often clinically mild or asymptomatic diseases (e.g., chlamydia, hepatitis C) are often incomplete because cases may not receive medical care or may not be diagnosed. The quality of the reported data also varies. Birth certificate data on birthweight, for example, are generally more reliable than some of the accompanying information, such as reports of birth defects or the mother's use of tobacco during pregnancy.

The significance of such limitations in these data depends on how the data are to be used. Estimation of reliable incidence and prevalence rates, for example, requires nearly complete reporting, whereas monitoring of trends depends more (within limits) on consistency of reporting than on completeness. For example, consistent and essentially complete reporting of births and deaths is the basis for calculation of comparable birth and death rates at the local, state, and national levels. In contrast, reportable disease data compiled at the national level are useful for monitoring disease trends even if they are not complete; however, these data are appropriate for more precise assessments of incidence rates only for those conditions for which reporting is essentially complete. And any variation in reporting practices from state to state means the resulting data will not be appropriate for assessing small differences in incidence rates across states.

Surveys

Surveys are an essential resource for population-based performance measurement data. Well-designed surveys produce information about an entire population by collecting data from a representative sample of that population. The population of interest in a survey is often defined by residence in a geographic area, such as a state or county, but may also be defined by other characteristics, such as age, place of employment, enrollment in a public assistance program (e.g., Medicaid), or use of a specific clinic. Continuing survey programs that have a defined schedule (e.g., the National Immunization Survey, the Behavioral Risk Factor Survey) can combine a stable core of questions, yielding results that can be compared over time, with changing sets of questions that can address topics of special interest. One-time surveys or surveys repeated on an irregular schedule have less value for performance measurement because they provide at

best a limited basis for comparisons over time. The use of surveys requires special expertise in such matters as questionnaire and sample design.

Surveys are particularly well suited to obtaining data for many measures of health status, functioning, and risk that depend on reports of behaviors, perceptions, and attitudes. They also are good tools for collecting information on general activities and events. Survey data are, however, vulnerable to misreporting and can be adversely affected by nonresponse. Respondents may misreport unintentionally because of recall errors or lack of knowledge (e.g., date of last illness or hypertension status), may refuse to answer certain questions, or may intentionally alter their responses on sensitive topics (e.g., drug use or even exercise habits). Careful questionnaire design can help reduce some forms of misreporting. Nonresponse is a concern because individuals who are missed may differ from the respondents in important ways (e.g., older or younger, lower or higher income, sicker or healthier) that cannot be determined with certainty. Despite such limitations, surveys may be the best or only option for obtaining data on key topics of interest.

The cost of surveys is a major constraint on their use. In contrast to data collection that occurs as a byproduct of other activities, such as restaurant inspections or health care visits, surveys require a set of specialized activities, including developing a sampling frame, selecting the sample, locating the eligible respondents, and gathering the survey data. For each of these activities, choices can be made that affect costs, but those choices may also affect the quality of the survey results. For example, telephone interviews tend to be less costly than in-person interviews, but cannot reach people who do not have a telephone.

Such cost tradeoffs should be weighed carefully. For some purposes, a factor such as telephone access may have little impact on the quality of the data, readily justifying the use of a less costly method of data collection. An analysis of National Health Interview Survey data, which were obtained through in-person interviews, found little difference in results between respondents who had telephone access and the overall responses, even when the analysis was restricted to persons below the poverty level (Anderson et al., 1998). Similarly, studies of the Behavioral Risk Factor Surveillance System suggest that its telephone-based methods are sufficiently reliable to justify continued use of this less expensive method (e.g., Arday et al., 1997). In contrast, a study focusing on health insurance coverage suggests that reliance on telephone interviews alone may not be adequate for some analyses (Strouse et al., 1997). It may, however, be possible to use baseline data from in-person interviews to adjust estimates based on data collected by telephone in subsequent rounds of a study.

Patient Records and Related Clinical Encounter Data

The detailed clinical records maintained by physicians, hospitals, health plans, and most other health care providers on each patient they treat are reposi-

ries for an array of data such as patient-reported health status and risk factors, clinical observations, diagnoses, procedures performed, medications prescribed, and results of laboratory tests. Access to clinical data from medical records would improve the analytic strength of many health survey and administrative data sets. However, these records have important limitations. Most patient records are still maintained in paper form, which makes it difficult to aggregate and analyze the data or integrate them into broader health data systems. Extracting data from paper records requires time-consuming and costly review of individual files. Research studies that require specific clinical data often review samples of records, but even that approach is likely to be too costly and time-consuming to be practical for the periodic reporting required for performance measurement. Furthermore, the completeness and consistency of records may differ across records or within a single record over time, and may vary more for certain types of information than for others. For example, numerical data, such as blood pressure readings, are more readily recorded in a consistent manner than are notes describing clinical observations.

There is widespread support for the development of computer-based patient records (CPRs), and considerable progress has been made in this area in recent years (Institute of Medicine, 1997). The CPR holds the promise that documentation of the process and outcomes of care will become a byproduct of the use of such an information system in the delivery of care, and that patient records will become a more practical source of data for performance measurement for both the health care industry and health agencies at the federal, state, and local levels. Major advances are needed in at least three areas, however, if more extensive use is to be made of clinical data in computerized form: standards defining the structure and content of electronic clinical records must be established, technology for converting natural medical language into standardized coding systems must be developed, and privacy concerns must be resolved.

Despite progress, there are still substantial differences and incompatibilities among the CPR systems now in use. Standards for the data elements included in patient records, the codes and vocabulary used to represent clinical data, and the format of electronic records are still evolving. Additional research and testing are also needed to move beyond prototype systems for converting natural medical language into medical procedure and diagnosis codes. Among the groups working on these CPR issues are federal agencies such as the National Library of Medicine and the Agency for Health Care Policy and Research, private organizations such as the Computerized Patient Record Institute, and various private companies. Progress toward a CPR should also result from the Health Insurance Portability and Accountability Act of 1996 (HIPAA) (Public Law 104-191), which directs the Secretary of Health and Human Services to promulgate guidelines for computerized medical records by August 2000. HIPAA also calls for establishing policies to protect the security and privacy of electronic health data transactions. Privacy concerns are an issue for all health-related data, but are

particularly acute for information contained in medical records. (Other provisions of HIPAA are reviewed elsewhere in this chapter.)

Computerization per se will not, however, overcome certain limitations inherent in patient records. For example, as used in most health care settings, patient records are not well suited to capturing information on patients' views about the care they receive. Clinical records can also be incomplete when people receive health care services from several sources, each of which maintains a separate record. For some of its performance measures, the Health Plan Employer Data and Information Set (HEDIS), version 3.0, compensates for such factors by requiring health plans to use data from a member survey rather than from administrative or medical records (National Committee for Quality Assurance, 1996; see also Chapter 2). For example, the rate of influenza vaccination among older adults is to be tracked with survey data because health plan members may receive these shots through community programs instead of their health plan.

Administrative Data

The operation of health programs typically generates substantial amounts of nonclinical administrative data that can be useful for performance measurement. Some of this information describes program resources or characteristics of program operation, such as numbers and qualifications of staff members or features of facilities used to provide services (e.g., number of laboratories meeting quality standards). Administrative records on population-based services can produce such information as the number and results of restaurant inspections, the number of immunizations administered at special immunization events, or the number of health education programs offered. Programs that provide services to specific individuals (e.g., substance abuse treatment or prenatal care) generate administrative records that contain information about those individuals and the services they receive. Administrative data produced by various other activities that are not specifically health related can also provide useful information for health programs. For example, traffic safety records can provide data on motor vehicle crashes resulting in injuries, and state corrections records can provide information on incarcerated adults with serious mental illness. In addition, administrative records can sometimes be used to identify a population for a separate survey-based data collection activity.

Most administrative data sets are created to serve operational purposes rather than the needs of performance measurement or other analytic tasks. Even so, they are a valuable resource with some advantages over other types of data. A recent assessment of the utility of administrative data for policy studies of public assistance programs provides useful insights for the health-related programs of interest to this panel (Joint Center for Poverty Research, 1998). Administrative data sets can offer detailed and generally accurate program information, large enough numbers of records to permit analyses of subgroups of participants, greater

state and local specificity and applicability than many national data sources, longitudinal information on programs and program participants, and low marginal costs for data collection.

At the same time, these data sets have important limitations for secondary uses such as performance monitoring. They generally cover a selected set of people and activities and are not necessarily representative of the population as a whole. In the case of health services, for example, such data sets have no information on individuals in a community who might be in need of those services but have not sought care. The data sets also may lack useful descriptive information on the economic and demographic characteristics of the individuals who are included. Measures such as program participation rates require that administrative data (the numerator) be supplemented by population data (the denominator) from another source. Information on outcomes and events that occur outside the framework of the program are rarely available. For example, the records of a substance abuse treatment program can produce data such as the number of participants who complete treatment, but will not directly capture the drug-related arrests of program drop-outs or the subsequent employment history of people who have successfully completed treatment. Similarly, records of a water treatment facility can provide data on observed levels of bacterial contaminants, but will not reflect outbreaks of waterborne disease. Linkages to other data sets (discussed below) can overcome some of these limitations, but the linkage process poses its own technical and policy challenges. These issues are discussed elsewhere in this chapter.

Use of program-specific data definitions can hinder or prevent valid comparisons across data sets. For example, one health program may define adolescents as young people between the ages of 12 and 18, while another may use ages 13 to 17. Greater coordination and collaboration and the development of standard measures may overcome some definitional differences. Yet other differences in data definitions reflect true variations in program features; if comparisons are necessary, those variations must be taken into account. Operational and design factors may also affect the usefulness of administrative data sets for purposes such as performance monitoring. Programs that serve families may not identify each family member separately, making it difficult to distinguish who received what services. If closed or inactive cases are dropped, the data set cannot provide a complete record of services or participants. And the installation of new or upgraded information systems (either equipment or programs) may result in lost or limited access to records created with the previous system.

Claims Data

A specialized administrative data resource that bridges the public and private sectors in health care is insurance and other third-party claims for payment for health services. An enormous quantity of data is produced from the billing and

payment of health insurance claims. In accordance with the administrative simplification provisions of HIPAA, standards for the format and content of electronic claims transactions are being established. Claims data have been used to study the effectiveness and outcomes of health care and may also have a place in performance measurement. As with the other data sources discussed in this section, however, their limitations must be kept in mind.

Claims data generally include only a minimal amount of clinical information (e.g., diagnosis, procedure performed) to document the fact that a covered service was provided and payment is owed. Moreover, medical conditions and treatments can be characterized in varying ways in insurance claims. This factor can reduce the consistency and comparability of claims records. Incentives such as higher reimbursement rates for certain types of care can encourage more deliberate changes over time in the content of claims data. The timeliness of these data can also be a concern. Greater use of electronic data interchange (EDI) allows faster claims processing, but delays of several months may still occur in filing and settling claims.

Another limitation of claims data is that they may not provide a complete record of services received by an individual or by a population in a given community or state because claims are submitted only for covered services and only for the individuals served by a specific insurer. Typically, a defined geographic area (a state or a community) is served by several insurers, each of which may offer many different insurance products that vary in scope and terms of coverage. In addition, Medicaid claims records may be managed separately by state agencies, and prepaid managed care plans generally do not generate claims records. With nearly universal participation in Medicare among those aged 65 and older, Medicare claims files have been more complete than other claims databases and therefore often more useful for state and local analyses. However, claims records are generally not available for Medicare services provided through prepaid managed care plans.

The experience of the State of Maryland in using Medicaid claims data in conjunction with public health initiatives illustrates both the strengths and limitations of such data when used for a purpose other than that for which they were originally collected (see Box 4-1). Although these data are a promising means of monitoring health care services for a vulnerable population, they do not capture all of the information that may be needed for some purposes.

Linkage of Data Sets

Data linkage involves matching records on specific individuals to other records for those individuals in the same or other data sets. This panel believes that in many cases, better performance measurement data could be obtained if selected data sets could be linked. As noted earlier, such linkages can overcome some of the limitations of specific data sets. This is especially true in efforts to

Box 4-1

Development of a Health Care Services Database in Maryland

In 1985, Maryland began developing person-based analytic files from Medicaid data. These files were used to conduct analyses that provided a basis for statewide public health initiatives such as diabetes management (Stuart, 1994) and the Maryland Access to Care program, which increased access to primary care for women and children participating in Medicaid (Stuart et al., 1990). The Medicaid data files were also used in research studies, alone or supplemented by data from medical records, to assess the quality or effectiveness of care (Starfield et al., 1994; Powe et al., 1996; Svikis et al., 1998). In another study, the claims records were used as a sampling frame for a survey that provided detailed information on a vulnerable population (Rubin et al., 1994).

The state's use of claims data was expanded in 1993 with the establishment of the Maryland Health Care Access and Cost Commission. The commission's tasks included the development of a medical care database to provide statewide information on health services rendered in all nonhospital settings, including health maintenance organizations and pharmacies. This mandate has been carried out using the computerized claims payment history files from third-party insurers in Maryland, including Medicaid, Medicare, and selected private payers (Stuart, 1995).

The commission's work has also highlighted the limitations of this type of database for public health applications such as tracking immunizations. In particular, it was anticipated that suitable data could be obtained from claims filed using standard billing forms (HCFA-1500 and UB-82/92). Private physicians, noting the requirement to complete multiple forms every time they immunized a child, pleaded for a unified system that required only one submission of data. Public health administrators pointed out that the HCFA-1500 form does not capture information on vaccine lot numbers, and standard procedure codes have not been established for administering some vaccines. In addition, since instructions for using the forms and codes for this purpose have not been standardized, recorded information could be subject to inaccurate interpretation (Stuart, 1995).

While the Maryland claims database cannot yet replace the public health immunization registry, the potential for using computerized claims files to reduce administrative costs and improve timely access to data that are useful for public health has been recognized.

relate health outcomes to services provided. For example, linking data from a community survey to administrative records from a prenatal care program could help identify eligible mothers who did not participate in the program and therefore do not appear in the administrative data system. Alternatively, program records could compensate for survey respondents' recall errors about numbers of visits or timing of specific services. Another approach to linking data sets is taken with some immunization registries: birth records are used to create an initial entry in the registry to which subsequent immunization reports are linked. In health care studies, efforts have been made to link multiple insurance claims for an individual to construct a more coherent picture of care for an episode of illness.

A particularly broad pilot project on data linkage that is relevant to the panel's earlier work on performance measures for emergency medical services was initiated by the National Highway Traffic Safety Administration (1996) of the U.S. Department of Transportation. The Crash Outcome Data Evaluation System (CODES), originally tested in seven states, is designed to link data on motor vehicle crashes, emergency medical services, emergency department care, hospital and outpatient care, rehabilitation and long-term care, death certificates, and insurance claims. Using these linked data, states have been able to explore such factors as populations at increased risk for injury (e.g., on the basis of age, alcohol use, or failure to use seatbelts), the consequences of specific types of crashes or injuries (e.g. collisions with pedestrians, abdominal versus head injuries), and the effects of delayed prehospital care.

Attempting to match records from separate data systems poses significant technical challenges. Reasonably successful techniques have been developed that rely on combinations of information, such as name, address, and date of birth, to establish highly probable matches. Use of unique personal identifiers might simplify the process of establishing exact matches, but such identifiers have not been uniformly employed. Provisions of HIPAA now call for adoption of these identifiers, especially for use in electronic health care data transactions, but there is serious concern that stronger privacy protections must be enacted before unique personal identifiers can be used with confidence or comfort (see National Committee on Vital and Health Statistics, 1997b). Even without the use of personal identifiers, the linkage of data sets must be undertaken only with firm assurance that personal privacy and the confidentiality of the data will be protected. (See the discussion of these issues later in this chapter.)

Steps Toward Integration of Data Sets

In the public sector, many states are working to enhance the integration and accessibility of health data (Mendelson and Salinsky, 1997; U.S. Department of Health and Human Services, 1998b). For example, Georgia has provided a single Internet access point to county- and state-level information from several data

sets, such as vital statistics and notifiable diseases. In Illinois, the Cornerstone system integrates client records for various maternal and child health services provided by local health agencies. Most states are also enhancing their Medicaid data systems and promoting the use of electronic claims transactions (Mendelson and Salinsky, 1997). The Center for Substance Abuse Treatment in the Substance Abuse and Mental Health Services Administration (SAMHSA) is working with three states to explore ways of linking state Medicaid, mental health, and substance abuse data sets to provide more complete information about clients and service use and to support the implementation of performance measurement.

Attempts to establish public-private partnerships to facilitate the integration of health care data have had limited success. For example, the Community Health Management Information System (CHMIS), proposed by the Hartford Foundation in the early 1990s, was envisioned as a community repository and resource for health care data for use in assessing the cost and quality of care offered in the community. Support for such a community-based approach, however, has been weakened by fundamental changes in the organization of health care services that have resulted in the growth of large regional and national insurers, integrated health care delivery systems, and managed care organizations (Starr, 1997). These organizations are now investing in their own information systems to provide corporate- rather than community-based analyses of cost and quality. Other obstacles included the technical complexity and expense of community-based systems, concerns regarding the confidentiality of patient records, and the reluctance of some health care organizations to share information with business competitors.

An alternative model, sometimes referred to as the Community Health Information Network (CHIN), has shifted the focus from the collection and storage of information for use in the community to the development of clearinghouses that would transmit information among diverse proprietary information systems maintained by insurers, managed care organizations, and individual hospitals and clinicians (Starr, 1997). Such efforts are hampered, however, by the proliferation of proprietary information systems using customized administrative transactions, which also impose a substantial burden on the health care system. For example, hospitals and physicians often find that each insurer and health plan uses a different claims form that requires somewhat different information. Administrative overhead is estimated to account for about 26 percent of health care expenses (U.S. Department of Health and Human Services, 1997c). The administrative simplification provisions of HIPAA are an effort to reduce this burden by establishing standards for electronic health data transactions. These standards can also be expected to facilitate the integration of administrative health care data into other applications.

ASSURING THE QUALITY OF DATA AND DATA ANALYSIS

Health professionals and policy makers seeking to use performance measurement in conjunction with publicly funded health programs must consider the quality, consistency, and comparability of the data available for this purpose and determine how to address the limitations of those data. The panel cannot offer simple, straightforward criteria for judging the quality of health data and data systems. However, because it is nearly impossible to evaluate data quality on the basis of summary measures, quality is an essential consideration at every step, from the planning of data systems and data collection to the calculation and use of final measures. These issues are not unique to performance measurement, and lessons learned in other contexts merit attention (see, e.g., Hoaglin et al., 1982; Bailar and Mosteller, 1992).

Ideally, the data used for performance measurement would be totally accurate and complete. In practice, however, data rarely meet these requirements. Many different factors may affect the quality of medical and scientific data, whether the data are used for scientific study, administrative purposes, or management oversight as in performance measurement. Some ways in which data can be compromised include inaccurate reporting, incomplete reporting, poorly designed survey samples, errors introduced during data processing procedures, inappropriate aggregation of detailed data to facilitate analysis, and inaccurate calculation of measures.

Standards for data quality and practices adopted to meet those standards should be based on informed assessments of the intended and anticipated uses of the data. Some consideration should be given to potential future uses of the data, but data systems should not be overdesigned in an effort to meet all possible but as yet unidentified requirements. Other concerns relate to the implications of data quality for analysis. By itself, a data set may appear to produce data of satisfactory quality. Problems may arise, however, if the characteristics of older data differ from those of newer data or if the data differ in important ways from other data with which they might be used.

Early and continuing advice from and participation by experts in such fields as statistics, epidemiology, and informatics can reduce the likelihood and severity of many problems involved in the design and use of data and data systems. For example, the use of observational and administrative data for performance measurement poses analytic challenges that differ from those for studies that can rely on more carefully controlled experimental data. Although opportunities to redesign the existing data systems that will provide much of the data for performance measurement will be limited, expert advice can help maintain or improve the quality of those systems. For any new data systems, expert advice early in the design of the system is particularly important because well-designed data systems can prevent many problems that are difficult, and sometimes impossible, to overcome by analytic techniques.

Policy makers who use performance measures should ensure that there is a review process to determine what problems are most likely to affect the data, what has been done to manage those problems, and (at least roughly) how large any residual problems are likely to be. Data and data systems should be held to high standards, but the use of reasonably good data with known limitations may be acceptable, even desirable, for some purposes given the opportunity costs of collecting better data. For example, data from a survey of teenagers, with all the biases inherent in such surveys, are likely to be better for determining the frequency of violation of laws restricting cigarette sales to minors than highly accurate and complete court records covering only violations that have come to judicial attention. Substantial investments of time and money in data that are not appropriate for the analysis at hand or in activities that will produce only marginal improvements in the data do not represent a good use of resources. In cases in which bias dominates random variation, for example, the benefit gained from stringent reductions in the random component of uncertainty (e.g., through use of larger samples) may not justify the cost involved.

Many observers agree that making data useful and important to those who produce the data creates a strong incentive for ensuring that the data are of high quality. Performance measurement may help provide such an incentive by requiring that data be used for internal purposes or for external reporting. As noted in Chapter 2, however, care must also be taken to avoid the creation of adverse incentives that could encourage deliberate distortions of the data to make performance measures appear more favorable than is warranted.

A few of the statistical and operational factors that can affect the quality of data and their analysis are reviewed briefly in the following subsections. These discussions provide only an introduction to potentially complex issues that should be addressed in more detail by those responsible for implementing performance measurement.

Random Variation and Bias

In collecting and using performance measurement data, policy makers and program staff must keep in mind the effects of random variation and bias. Some degree of random variation should be expected among otherwise similar measurements. For example, small year-to-year changes can occur in the number of infant deaths without representing a meaningful change in the underlying infant mortality rate. Similarly, two independent random samples drawn from the same population are likely to produce slightly different but still representative estimates of the average age of all of the members of that population or of measures such as the percentage of adults who have had their blood pressure checked in the past 2 years. The effect of random variation tends to be greater in measures based on small numbers of events or small sample sizes in surveys. For example, random variation in the annual number of infant deaths will have less impact on

the stability of the national infant mortality rate than on the rate in an individual community. Statistical techniques can be used to estimate the likely contribution of random variability to a sample estimate or to the difference between two measurements (e.g., infant mortality rates for successive years). In some situations, the variation arising from small numbers of cases can be reduced by using measures that pool related but individually rare events. For example, a community might measure the percentage of the population using any illegal drugs rather than attempting to measure separately the use of several different drugs.

Bias reflects systematic distortions in the data and poses a more serious challenge to successful use of those data. Bias can be introduced in the design of data collection procedures and in the data collection process itself. In surveys, for example, bias may result from a sample design that excludes certain groups (e.g., households without telephones, as discussed earlier). Bias may also result from differing response rates by specific population groups (e.g., fewer responses by single adults than by married adults with children) or from intentional misreporting (e.g., underreporting of tobacco and alcohol use in panel surveys). Data from registries and administrative records can be affected by systematic differences in the populations they cover or by incomplete or inaccurate reporting. For example, population differences could be reflected in health insurance claims. A firm with an older workforce would tend to have more claims related to care for such conditions as diabetes and hypertension than a firm with younger employees. Financial incentives associated with variations in reimbursement rates may also influence the way diagnoses or health services are characterized in health insurance claims.

Although bias is undesirable, it may not make data unusable. When bias is constant (over time, over a geographic location, over a population segment), it cancels out of many kinds of analysis. For some purposes, imperfect data with a constant bias may be more useful than continually improved data. For example, if the incidence of a disease has been underreported by a consistent 20 percent over a period of years, the trend in its incidence can still be assessed, and any ratios and proportions calculated using data from that period will be accurate.

Data Management

Other factors that can impair data quality include coverage problems that can occur at the time the data are collected or during the data processing phase. Records may be duplicated, inappropriate records may be included, or appropriate records may be missing. Duplication of records may occur if multiple reports about a single individual are received from separate sources and cannot be matched. Some states, for example, require that health care providers and laboratories submit reports on cases of HIV infection without the use of individuals' names, but difficulty can be encountered in matching the reports on a given individual that come from separate sources. Data linkage can also cause problems if

records are matched incorrectly. Moreover, data sources such as registries and administrative data systems can be affected by delays in receiving or entering records that result in missing records at the time a report is produced. A data set may also be incomplete because of such factors as a very low response rate among those who believe that a survey may harm their interests or failure to identify and collect the death certificates of all cancer patients in a given cancer register.

Other problems arise if data are used incorrectly to construct performance measures or measures that may be used for other operational or policy purposes. For example, a measure of immunization rates among 2-year-olds will be flawed if either the numerator or the denominator includes children of the wrong age. Audits of health plan performance measures by the National Committee for Quality Assurance (1997) found average error rates of 20 percent, denominator error rates of up to 63 percent, and numerator error rates as high as 72 percent.

Attention to operational policies and practices throughout the collection and processing of data to produce performance measures is likely to help ensure that performance measurement is based on high-quality data. The National Committee for Quality Assurance (1997) has specifically recommended that health plans implement routine data-quality audits to improve the accuracy and completeness of their clinical and administrative data sets. Among the considerations highlighted is good documentation for all steps involved in data collection and processing, including both instructions for each step and records of what was actually done. Audits can verify the accuracy of individual data elements and of the measures calculated using those data. Automated edit checks can test the consistency of data entries. For example, a record showing a 10-year-old respondent in a survey of adults can be flagged for review. Automated and manual checks at the data processing stage should ensure that data are being drawn from appropriate sources (e.g., survey data for the correct year), that calculations are being performed correctly, and that the data being used are consistent with established definitions. For example, if adolescents are defined as 14 to 17 years of age for a given measure, data for those aged 13 to 17 do not meet this definition.

Challenges in Data Analysis

As discussed in this report, performance measurement is most likely to rely on agreed-upon measures that are widely accepted as representing specific programmatic activities and that use data from existing, defined sources. Once those measures have been selected and the data produced, policy makers and other analysts may face several challenges in the successful use and interpretation of the data.

Almost every statistical analysis requires some sort of a statistical model to summarize the data and guide interpretations. A correct model can add great strength to the analysis, but an incorrect model can lead to unreliable findings. Because the correct model is generally unknown, experienced analysts may make

a critical contribution through their inferences about the form of an appropriate model. For performance measurement, this observation relates back to our understanding of the evidence that links processes and capacity to health outcomes. If that understanding is good, it becomes possible to select measures of outcomes, processes, and capacity that provide reliable signals of progress toward health goals. If, however, that understanding is incomplete or flawed, the process and capacity measures selected may provide little insight into the change (or lack of change) in health outcomes.

The issue of data comparability noted earlier represents an important challenge to data analysis. Comparisons among groups or over time are of particular relevance for performance measurement. The panel anticipates that performance data used in the framework of performance partnership agreements will frequently become the basis for comparisons across states. Similarly, states may make comparisons across counties, cities, or other community units. Data comparability is also a common issue in the interpretation of changes over time. Results can be affected, for example, by differences in methods of collecting the data, in the health care or program environments, and in the underlying characteristics of the populations being measured. A study designed specifically to test the ability of five states to report comparable data for a set of mental health performance measures demonstrated that such differences are currently an obstacle for performance measurement (National Association of State Mental Health Program Directors Research Institute, 1998; see also the discussion of this study in Chapter 3). Even efforts to improve existing data systems (e.g., through more complete coverage or better questionnaire design) have the undesirable, though often acceptable, side effect of hindering the interpretation of time trends. Lack of complete comparability does not preclude the use of the data, but it necessarily affects the nature and strength of the conclusions drawn from analysis of the data. Several common concerns related to data comparability are discussed below.

Concerns Related to Differences in Data Collection Methods

Different methods of collecting data regarding a particular phenomenon can produce different findings. For example, within the broad domain of surveys, the specific techniques employed will affect the accuracy of the data obtained and therefore the comparability of those data. Substance abuse rates ascertained from computer-aided interviews may be more accurate than those derived from telephone or in-person interviews (see Turner et al., 1998). Similarly, surveys that rely on self-reporting may produce more accurate data on respondents than surveys that allow reports about an individual by another person (i.e., proxy responses), although the use of proxies may improve data quality by reducing the nonresponse rate.

A different circumstance is illustrated by estimates of current tobacco use based on tax records as compared with estimates based on survey reports. Tax

records nearly always show substantially higher use than do survey data, apparently because tobacco use is underreported in surveys. Although the tax records themselves may be less than perfect, it appears likely that a substantial fraction of the tobacco sold and taxed will be used. Unlike surveys, however, tax data cannot provide information on the characteristics of those who purchase tobacco products. The findings from these two data sources might be used together to develop an adjustment factor for inflating the survey data on tobacco consumption to match estimates from tax receipts.

When data sources are as fundamentally different as administrative records and surveys (as in the example just cited), analysts may be more alert to the hazards of direct comparisons than when the differences are less obvious. If, however, a major discrepancy in findings between two or more data sources cannot reasonably be accounted for by differences in the methods of collecting or analyzing the data, analysts must consider which, if either, source should be used. If this choice is not clear, new data collection efforts may be warranted.

Concerns Related to Differences in the Program Environment

External factors in the program environment can affect the results of performance measurement in ways that are unrelated to program activities or goals and should be considered in interpreting performance results over time or across groups. Special circumstances, such as natural disasters or unrelated disease outbreaks (e.g., unusually high rates of influenza), might affect performance measurement results through either a deterioration in health outcomes or a reduction in the resources available for program activities. Apparent rates of disease incidence may also be affected by such factors as increased awareness of a given health problem or new health care technologies that alter patterns of detection and treatment of disease. For example, reported incidence rates for prostate cancer increased from 79.8 per 100,000 in 1980 to 132.0 per 100,000 in 1990 (both age-adjusted to the U.S. population in 1970) (Ries et al., 1997). During this period, increasingly widespread use of the prostate-specific antigen test led to a marked rise in the number of diagnosed prostate cancers without evidence that the underlying incidence of the disease had changed substantially and without a corresponding change in reported mortality rates. This change is readily attributable to a specific factor, but the factors underlying other differences can be less obvious or entirely hidden. Sorting out such matters generally requires the application of both statistical and subject-matter expertise.

Concerns Related to Differences in the Characteristics of Populations

Health outcomes are often closely linked to biological and social risk factors. Since the nature and distribution of these factors can be expected to vary across the populations being served in various programs and geographic areas, some of

the differences in outcomes may be attributable to these variations rather than to true differences in program performance. Methods are needed to adjust performance data for differences in important covariables over time or between comparison populations (see, e.g., Rothman, 1986; Gordis, 1996). Without adjustment, comparisons may often be difficult to interpret. For example, apparent variations in performance might be a reflection of differences in the characteristics of program participants (e.g., educational attainment, access to transportation), the general population in a state or community (e.g., average age), or the socio-economic and other characteristics of the states or communities served (e.g., unemployment rates, population density).

One method of accounting for such differences is stratification—the calculation of performance measurements separately for specific population subgroups (e.g., younger and older age groups)—which will provide more comparable results within those subgroups. This approach may not be feasible, however, if available data sources do not identify the subgroups of interest or if small numbers of cases compromise the reliability of the subgroup measures.

Differences in the mix of subgroup characteristics across populations can also be addressed by adjustment methods that permit calculation of a single measurement for each population to be compared. The “direct” adjustment method is one of the most widely used. With this method, population-wide rates are calculated by applying the observed subgroup measurements from each population of interest (e.g., age-specific rates for smoking or completion of substance abuse treatment) to the equivalent subgroups (e.g., age groups) in a single “standard” population. For example, in comparisons of cancer incidence, which is generally higher in older age groups, rates are often “age adjusted” using this method to ensure that observed differences can be attributed to disease incidence (or its detection), rather the age distributions of the populations.

More complex forms of statistical analysis offer other ways of accounting for differences in population characteristics. For example, “risk adjustment” techniques have been used to account for differences in initial severity of illness among patients in comparisons of clinical outcomes, such as hospital mortality rates for cardiac surgery (see, e.g., Luft and Romano, 1993; Iezzoni, 1994; Landon et al., 1996). These techniques are, however, still evolving, and different severity-adjustment methods have been shown to produce differing performance results (Iezzoni, 1997b).

In theory, similar adjustments could be made in evaluating performance data for health programs, but specific methods of adjustment have not yet been adopted for this purpose. Doing so would require determining which factors are appropriate to use in an adjustment, developing the statistical model to be used, and ensuring the availability of the necessary data on the adjustment factors.

Any adjustment method must, of course, be used carefully. One concern is ensuring that adjustment does not disguise meaningful differences in program performance among subgroups in the population. An adjustment based on

income, for example, might mask different levels of performance for lower- and higher-income groups. Another concern is that currently limited knowledge regarding the relationships between health outcomes and many social or biological factors may lead to inappropriate uses of adjustment. Determining which factors provide an appropriate basis for adjustment of performance data will require careful consideration of both the technical and policy implications of these methods.

The Drug Evaluation Network Study illustrates attempts to make such adjustments, in this case for comparisons of drug treatment centers.¹ The Addiction Severity Index (ASI) (see McLellan et al., 1992), an extensive patient interview instrument, is being used to collect information about the characteristics of the substance abuse patients served, including employment, legal involvement, family, and psychiatric problems, as well as the nature and extent of their illegal drug use.

Implications for Data Analysis

What are the implications of the preceding observations for data analysis and for policy that flows from that analysis? The experienced analyst is far from helpless even in the face of serious bias and/or incompatibilities in the data. First, however, the analyst must be familiar with the details of the data collection methods, as well as the procedures used to prepare the data for analysis and the specific format of the resulting data files. Every increment in understanding may reveal additional influences on the data that should be considered. Second, analysts and policy makers need to know that bias is likely to be critically important, but that there are some means to control or understand its influence. Third, ordinary tests of statistical significance and confidence bounds do not capture the broad and perverse effects of bias, and hence may be seriously misleading. Fourth, analysts and policy makers must anticipate that different methods of answering a question will sometimes produce apparently incompatible results. Similar uncertainties about the meaning of results arise even when only one set of observations has been made; if the data had been obtained by other reasonable methods, the results would have been different. Thus the policy analyst and the statistical analyst must work together to understand the strengths and limitations of a specific data set so that policy will be robust against problems in the data as obtained.

¹The Drug Evaluation Network Study and its use of the Addiction Severity Index were described to the panel by Thomas McLellan (professor of psychiatry at the University of Pennsylvania and senior scientist at the Veterans Administration Center for Studies of Addiction at the University of Pennsylvania) and colleagues at a workshop held by the panel in July 1997.

DEVELOPING AND IMPLEMENTING STANDARDS FOR DATA AND DATA SYSTEMS

One of the difficulties the panel faced in the first phase of its work was the limited availability of data that are comparable across states for use in performance measurement. Achieving greater comparability will require more standardization in the content and methods of data collection, in the coding and vocabulary used to record data, and in the selection and definition of performance measures. In addition to standards for the substantive comparability of data, effective use of information technologies requires standards for the format in which data are stored and transmitted. These issues are relevant for all forms of health data, from infectious disease and vital records reporting, to survey data, to clinical and administrative records.

Standardization has proven elusive for at least two reasons. First, health data are often complex. Second, many data systems have been developed independently to meet local needs, and it can be difficult to reach consensus on standards that may require substantial change in those data systems or may seem less likely to meet those local needs. With regard to choosing performance measures, lack of consensus can also reflect a field's need for further development of a framework for assessing performance. Although standards can be imposed through regulation and legislation, the panel favors a collaborative approach based on the participation of interested parties at the national, state, and local levels to ensure consideration of a broad range of views.

Standardization Activities

Although much remains to be done to improve standardization in methods of data collection, in the coding of health data, in the formats for storing and transmitting data, and in the definition of performance measures, many activities in the public and private sectors are making useful contributions in these areas. Several of these efforts are briefly reviewed below.

Centralized Data Systems

Centralized data collection efforts at the national level (sponsored by the federal government or organizations with national and multistate agendas) can use comparable definitions, questions, and methods across many or all states. Most of these activities result in national-level data, but usually cannot provide subnational estimates. One exception is the National Immunization Survey, a random-digit-dial telephone survey that yields state and regional estimates of immunization rates for children aged 19-35 months. This federally run survey uses comparable data collection methods across all states and regions, and comparisons of rates of immunization can reasonably be made among states.

Federal-State Collaboration in Public Health

Often, a centralized approach to data collection and analysis is too costly and inflexible to provide adequate state- and local-level detail, and other means of achieving comparability are necessary. Collaboration by the states and the federal government has led to a few well-recognized successes in harmonizing independent state systems. A notable success is the national vital statistics system, a cooperative state-federal program through which recommended forms and procedures for the collection and reporting of vital records data have been developed. Data collected by each state vital records system are reported to and compiled by the National Center for Health Statistics to produce national totals. The National Notifiable Diseases System, which relies on state reporting of new cases of specific conditions, was enhanced in 1990 by the development of standard case definitions for nationally reportable conditions (Centers for Disease Control and Prevention, 1997). SAMHSA compiles the Treatment Episode Data Set from a minimum set of data collected by states on clients admitted to substance abuse treatment programs that receive funding through the state substance abuse agency.

Under the Behavioral Risk Factor Surveillance System (BRFSS), the Centers for Disease Control and Prevention (CDC) has worked with the states to reach agreement on a core set of questions and standard sets of supplemental modules. CDC provides overall support and technical oversight, but individual states administer the survey and have the opportunity to add their own questions. However, because sampling design and data collection methods vary among states, comparisons of BRFSS data among states must be made cautiously, and it has not been possible to aggregate state estimates into national totals. State surveys may, for example, have significantly different response rates, and users of the data should consider how nonresponse bias may have affected the estimates.

Healthy People

Over the past 20 years the Healthy People initiative has provided a framework for establishing a common, national set of measurable health promotion and disease prevention objectives (U.S. Public Health Service, 1979; U.S. Department of Health and Human Services, 1991), and most states report using Healthy People 2000, at least in part, to guide the development of similar state-level health objectives (Public Health Foundation, 1998). The national objectives for 1990 and 2000 were not created with performance measurement in mind, but initial proposals for Healthy People 2010 specifically call for efforts to link the objectives to performance measurement activities under the Government Performance and Results Act (GPRA) (U.S. Department of Health and Human Services, 1998a).

Healthy People is contributing to standards that will be useful to performance measurement by promoting the adoption of specific measures for tracking

progress toward identified objectives and the development of detailed operational definitions for those measures. Some of the Healthy People 2000 measures have been adopted as performance measures for federal block grants (e.g., the Maternal and Child Health Services Block Grant [see Chapter 2] and the Preventive Health and Health Services Block Grant). Also, several of the measures proposed by this panel in its first report (National Research Council, 1997a) are quite similar to Healthy People 2000 measures, allowing for differences in data sources for state-versus national-level data. A series of reports from the National Center for Health Statistics is providing detailed specifications for operational definitions and data sources for the measures used to track progress toward the national Healthy People objectives (e.g., Seitz and Jonas, 1998). This information can help states and communities employ comparably defined measures.

Mental Health and Substance Abuse

Publicly funded programs in mental health and substance abuse are closely involved with the delivery of personal health services, often through providers in the private sector (see Chapter 3). A variety of data collection activities have developed, many of which have focused on services and service providers. The federally initiated Mental Health Statistics Improvement Program (MHSIP), for example, has been an essential resource supporting the development of information systems for public mental health services. Currently, however, comparable state-level data are limited for both mental health and substance abuse programs. As interest in treatment outcomes and performance measures has grown, state mental health and substance abuse programs have recognized the need to develop new and more comparable measures.

In 1997, the members of the National Association of State Mental Health Program Directors (NASMHPD) (1998) adopted a standardized performance indicator framework (see also Chapter 3). Using this framework as a guide, NASMHPD is working closely with SAMHSA and other organizations to identify and test measures that all states can use. Similarly, the National Association of State Alcohol and Drug Abuse Directors is working with its state members and SAMHSA to achieve consensus on a framework for performance measurement for substance abuse treatment and to develop detailed specifications for measures. Work is also being done on standard measures for substance abuse prevention.

Performance Measures in Health Care

As was noted in Chapter 2, several organizations are actively involved in the development and use of performance measures in health care, most notably the American Medical Accreditation Program, the Foundation for Accountability, the Joint Commission on Accreditation of Healthcare Organizations, and the National Committee for Quality Assurance (NCQA). Standard measures, data

definitions, and data collection systems are being developed for assessing providers, health care facilities, and health plans to determine whether accreditation requirements are being met and to provide purchasers and consumers of health services with comparative performance information. As these performance measurement programs are implemented, they will tend to encourage greater consistency in the health data components related to those measures. NCQA (1997) has also emphasized the importance of adopting recognized standards for the structure and content of both clinical and administrative components of health plan information systems.

Standards in Health Informatics

The increasing computerization of health data and the proliferation of incompatible information systems have generated efforts on many fronts to standardize various elements of the structure, function, and content of these information systems and of the format for transmitting information among systems. Organizations such as the American National Standards Institute (ANSI), the American Society for Testing and Materials (ASTM), Health Level Seven (HL7), the Institute of Electrical and Electronics Engineers (IEEE), and the National Uniform Claims Committee (NUCC) serve as private-sector forums for voluntary collaboration among parties interested in formulating standards for specific information system features.

Other groups are focusing on the development of standard coding sets and vocabularies for recording clinical information such as symptoms, diagnoses, procedures, and laboratory findings; examples of these coding sets and vocabularies are the International Classification of Diseases (ICD), the Current Procedural Terminology (CPT), and the Systematized Nomenclature of Medicine (SNOMED). No one coding system or vocabulary has become a comprehensive standard, and the Unified Medical Language System (UMLS), a project of the National Library of Medicine (1998), provides a “translation” tool to link information represented using these varying systems.

Health Insurance Portability and Accountability Act of 1996

HIPAA should result in substantial advances in the standardization of health care data and data systems. The administrative simplification provisions of HIPAA direct the Secretary of Health and Human Services (DHHS) to adopt standards for electronic transmission of administrative and financial health care data (see Box 4-2); for unique health identification numbers for health plans, health care providers, employers, and individuals; for code sets for data elements used in health care transactions; and for security of electronic transactions. The

Box 4-2
Administrative and Financial Transactions Covered
Under Standardization Requirements of the
Health Insurance Portability and Accountability Act

- Claims or equivalent encounter information
- Coordination of benefits information
- Referral certification and authorization
- Enrollment and disenrollment in a health plan
- Eligibility for a health plan
- Health care payment and remittance advice
- Health plan premium payments
- First report of injury (for workers' compensation)
- Health claims status
- Health claims attachments

act also directs the secretary to promulgate guidelines for computerized medical records within 4 years. DHHS will base HIPAA standards on existing standards in any of these areas that have been developed, adopted, or modified by standards-setting organizations accredited by ANSI.

The use of electronic transactions is not required, but if they are used, they must adhere to the HIPAA standards. Despite anticipated high initial costs of modifying or developing information systems to implement the new transactions, overall savings are expected to amount to billions of dollars (Office of the Secretary, U.S. Department of Health and Human Services, 1998). Standardized enrollment, coding, and billing formats will eliminate the need for health care providers to customize transactions to the varied requirements of many different health plans and insurers.

The impact of HIPAA will extend to state and local health departments and other health-related agencies. Those that function either as payors or as service providers seeking reimbursement will have to implement information systems that use the transaction standards. Standardization of data elements, data definitions, transaction formats, and code sets should aid the conversion of health encounter data into public health data. For example, with a standardized transaction format and standardized electronic data interchange, it should be possible to piggy-back notices of reportable illness on an electronic transaction. This process offers the potential for greatly enhancing the quality and timeliness of these reports. In addition, because a separate report would no longer be needed, the percentage of cases that are reported could be expected to increase.

Balancing Standardization and Change

As discussed earlier, greater standardization of data, data collection methods, and measures is essential to permit comparisons of performance over time or across groups. However, this standardization must be pursued thoughtfully. The participants in the performance measurement process must have an opportunity to gain experience with both the conceptual and practical aspects of performance measurement, and the process must be seen as a continuing activity that allows for the reassessment and revision of standard measures. In its first report the panel emphasized that the proposed measures of health outcomes and risk status were reasonable candidates for use in most states (if the necessary data were available), but that process and capacity measures had to be selected to match the particular program strategies that a state had adopted. The panel also observed that greater consensus regarding appropriate measurement domains may sometimes be necessary before performance measures can be proposed.

There is a risk that setting standards for performance measures or data sources will discourage improvement and innovation. Current limitations on the availability of data could, for example, encourage adoption of “least common denominator” measures for which data are widely available, rather than better measures for which new data collection efforts would be required. The desire for continuity of measures over time could also work to discourage constructive changes in the selection and definition of performance measures or in data collection and analysis.

Policy makers and others who develop and use performance measures need to recognize that they must be engaged in a *continuing* process in which measures and data are reviewed and revised in response to advances in knowledge and changes in program practices and priorities. Within a framework that strives for comparability, this process should allow for the introduction of new measures, the acceptance of new data sources, and the adoption of new techniques for data collection and analysis. As new measures and data systems are introduced, efforts should be made to calibrate them against previously established data systems to facilitate continued use of the data generated by those older systems in longitudinal analyses. The review process should also ensure that the measures and data sources already in use continue to be suitable and are used in appropriate ways. For example, for the most recent revision of the HEDIS measures (version 3.0), NCQA (1996) instituted a supplemental “testing set” of measures to be examined further before being adopted as required measures. With HEDIS 3.0, NCQA also established a standing Committee on Performance Measurement to oversee an ongoing process of reviewing and testing measures and to develop a research agenda for the development of new measures.

For a review process to be credible and acceptable, those whose performance is to be measured must have an effective means of participating in the deliberations conducted to select and review performance measures. For the public-sector

programs that have been the focus of the panel's attention, formal mechanisms should be developed to ensure that each major group of stakeholders at the state and local levels is a full partner in such discussions. The regional meetings organized in response to the PPG proposal were a welcome opportunity for broad participation, but are not a viable model for a continuing forum for discussion of program priorities, performance measures, or data resources. Federal, state, and local governments should ensure that policy, program, and technical perspectives are all represented, and might work with various organizations to identify representative participants from these constituencies for such an effort. Examples of these organizations include the Association of State and Territorial Health Officials, the National Association of County and City Health Officials, the Council of State and Territorial Epidemiologists, the Association of Maternal and Child Health Programs, the National Association of State Mental Health Program Directors, the National Association of State Alcohol and Drug Abuse Directors, the National Association of Local Boards of Health, and the Association of Public Health Laboratories. (Box 5-1 in Chapter 5 lists additional organizations that might be involved in these activities.)

ENHANCING PERFORMANCE MEASUREMENT THROUGH ADVANCES IN INFORMATION TECHNOLOGY

Advances in information technology are changing the environment not only for performance measurement, but also for many aspects of the health data infrastructure that support decision making for policies and programs (see, e.g., Lasker et al., 1995). These developments include improved capabilities for linking and merging electronic data sets, access to enhanced analytic resources as smaller computers become capable of using more powerful software to analyze larger data sets, and vastly expanded desktop access to information and options for data collection and transmission through the Internet and the World Wide Web. These advances in information technology enhance the ability to monitor the performance of state and local health agencies, as well as private providers of health care services. Furthermore, these developments can improve the availability of health performance information through new methods of communicating performance results to key audiences. There is, however, great variability among states and communities in access to and expertise in the use of information technologies. Noted here are a few of the developments in information technology that should enhance the ability to implement performance measurement.

Data Collection and Transmission Technologies

Technology is providing new options for data collection and transmission that can improve the quality and timeliness of the data. The widespread adoption of electronic birth certificates, for example, allows hospitals to enter birth certifi-

cate data directly into an information system that can check for missing or inconsistent data and then transmit the record to the appropriate office to register the birth. This eliminates the need with the older, paper-based reporting process for hospitals to forward written records that must then be transcribed in a central office, and for that office to send back to the hospital (sometimes more than once) questions about missing or suspect data that must be resolved before the birth record can be completed. For the most part, electronic birth certification is currently being used to automate the paper record process, but it could become the core of a more comprehensive information system on infant health that would link birth certificates with other data sources, such as records on prenatal care, metabolic disease screening, and immunization (Starr and Starr, 1995).

The Drug Evaluation Network Study, mentioned earlier, also illustrates new technological capabilities in data collection and analysis. Trained staff record information collected during an extensive patient interview directly into laptop computers. This allows the data to be monitored for invalid or inconsistent entries during the course of the interview and transmitted electronically to the researchers conducting the study. The electronic linkage between the treatment centers and the study staff makes it possible to update the interview protocol overnight to address policy changes or specific concerns about the nature of illegal drug use across the country.

Data Management and Analysis

Data management systems such as relational databases and data warehousing make it possible to maintain data in many separate files and link data from those files as needed. These systems can store information that includes personal identifiers, or they can be based on anonymous data records for which identifying information has been replaced by system-specific codes that allow records to be linked, but do not relate them to identifiable individuals.

The addition of geographic detail to health records and other types of information, often referred to as geocoding, can enhance the analytic value of many kinds of health data. Geocoded data can be grouped into geographic subunits for analysis. For example, responses from a state's Behavioral Risk Factor Survey might be grouped by county or other substate region to gain additional insight into possible differences in risk behaviors and program impacts or needs across the state. It may also be useful to include geographic information, such as the distance between a substance abuse client's residence and treatment site, in analyses of program outcomes. For some purposes, specific addresses may be needed, but even zip codes can provide useful geographic information for many analyses.

Geocoding also makes it possible to use new mapping technologies to display and analyze data. These geographic information systems (GIS) can capture and plot data from multiple sources to examine the spatial relationships among several factors that may affect health outcomes or health program services (see,

e.g., Clarke et al., 1996). For example, data on birth outcomes (e.g., birth weight, prematurity) might be plotted with data on the residence of the Medicaid-eligible population and the location of such services as health care providers, child care facilities, and grocery stores. The capture of geographic data for GIS is being enhanced by data collection systems that can record specific geographic coordinates by using satellite-based global positioning systems (GPS).

Computer-Based Patient Records

In health care settings, the use of information technology is expanding beyond the management of administrative and financial data to computer-based clinical information systems. As discussed earlier, there is great potential for CPRs to meet the need for timely and accurate clinical information that is difficult to access with traditional paper records (Institute of Medicine, 1997). Prototype CPR systems can convert natural medical language into medical procedure and diagnosis codes. Work is also being done to integrate decision support tools into these systems by incorporating clinical knowledge resources such as accepted treatment protocols. Based on these protocols, deviations or oversights in patient management can be detected and alternatives suggested.

Individual institutions have made progress in developing CPR systems. However, further advances will require not only technological innovation but also organizational and policy changes, such as greater consensus on clinical vocabularies and greater acceptance of changes in methods of recording information by those who use such systems. For example, voice-activated interfaces and increased familiarity with computer use are overcoming the past resistance of many physicians to typing notes directly into the patient record. The high cost of current CPR systems is another barrier to their wide acceptance. Further refinement of the technology and cost reductions are key factors for more widespread adoption.

The Internet and the World Wide Web

Both health data systems and performance measurement are affected by the extraordinary growth of the Internet and the World Wide Web as tools that facilitate communication and data exchange. Performance data can be collected and submitted to central repositories for processing and analysis and then made available to a broad range of interested parties. It is even possible to provide access to data in a form that permits customized analysis (see Box 4-3). Although use of these tools has expanded rapidly, they may not yet be appropriate for certain communities. Some of the more rural parts of the country may still lack affordable access to the high-speed telecommunications services that substantially enhance the utility of the Internet and the World Wide Web.

Box 4-3

Using the Internet and the World Wide Web

SHARING CURRENT INFORMATION

Recent Wisconsin legislation requires each community to conduct a biannual assessment of its population health needs and to develop a plan that reflects community health priorities. The scope of the assessment requires information from more than 20 different state data systems. To facilitate the process, the state's Department of Health and Family Services agreed to provide information to each community in an easily accessible format. Because of the recurrent nature of this requirement, the state needed to automate the process. A system was constructed that dynamically builds a narrative and graphic community-specific report card. This process involves five steps:

1. The PC-based system uses a statistical package to extract the desired data from the various state data systems.
2. The system transfers the data into a standard spreadsheet, which dynamically builds the agreed-upon outcome measures.
3. The measures are passed to a graphics program that creates a series of comparative charts.
4. The charts from the graphics program are loaded automatically into a preformatted narrative report.
5. The narrative report is published on the World Wide Web (and in hard copy) for universal access.

The state's communities can now develop plans using an integrated data report. When the data sets are updated, 89 community-specific reports are generated in less than 2 hours.

SOURCE: See <<http://www.dhfs.state.wi.us/AIM/index.htm>>.

RESPONDING TO AD HOC REQUESTS

Demands for state health data by internal and external users were overwhelming data agencies in Utah. The state is now attempting to meet these needs by providing access to data via the Internet. The Utah

system gives users access to linked data sets from which personal identifiers have been removed. A point-and-click interface executes a statistical package that generates the requested data. Results based on small numbers of cases are automatically suppressed. The system minimizes the demands posed by routine requests, so that data processing staff can focus on designing new applications. The ad hoc query component of this system also serves a similar function in an Intranet arrangement that provides previously authorized users access to more detailed information than that available to the general public.

SOURCE: Utah Department of Health, Office of Health Data Analysis; see <<http://hlunix.hl.state.ut.us/hda/>>.

SUBMITTING AND CHECKING DATA

Monitoring the long-term care received by nursing home residents is a long-standing requirement in the federal-state Medicaid partnership. A new emphasis on assessing health outcomes requires the collection of clinical data using the Nursing Home Minimum Data Set. Residents are assessed a varying number of times, and their Medicaid status (whether on Medicaid or not) is reported to the state. Manual reentry of this information is not feasible given the volume of reports. Several states and the Health Care Financing Administration worked with the nursing home industry to develop an Internet-based reporting system that encrypts residents' records. In the Wisconsin implementation, a nursing home records the required data in an in-house record system or directly enters the information using an Internet application. The data are submitted to the state electronically using a standard web browser and proceed automatically through state and federal edit checks. Inconsistencies and gaps are reported directly to the nursing home for correction. The state receives a problem log used to monitor volume and administrative issues. Concurrently, the system updates the various databases on nursing home residents. Standardized performance reports are generated for the state surveillance team to help focus their inspections. Both required and experimental measures are available in the system. Interim reports are also available to the nursing home to facilitate early intervention aimed at improving performance.

SOURCE: Panel member interviews with agency staff.

Limits of Technology

Developing and maintaining information systems designed to be long-lived is challenging when the technology is evolving at a rapid pace. As equipment and software advance, an agency may lose easy access to data from either obsolete systems or systems more advanced than its own. Incompatibilities can emerge in equipment, storage media, and programs, especially if information systems are developed and maintained in isolation. Collaborative efforts to develop information system standards can foster the evolution of independent information systems able to exchange information successfully. It is also essential for policy makers and data system managers and users to ensure that information technologies are employed only in ways that maintain the confidentiality of health data and protect the privacy of the individuals to whom the information applies. This issue is discussed in the next section.

PRIVACY, CONFIDENTIALITY, AND SECURITY OF HEALTH DATA

Protecting the privacy, confidentiality, and security of all forms of health-related data is a critical consideration in the collection and use of data for performance measurement. The public is concerned that personal health data may be used in detrimental ways, particularly as information technologies become more powerful and more pervasive. Fears that disclosures such as HIV test results or records of mental health or substance abuse treatment could lead to loss of employment or refusal of insurance may be especially acute. This concern is creating pressures for stricter technical and policy controls on access to and use of health data. At the same time, health policy makers and researchers worry that overly strict controls may hinder responsible uses of the data for research, performance measurement, and other such purposes aimed at controlling health threats or improving health and health services. Linkage of data sets can be a source of special concern because combining data in ways that may not have been anticipated when personal information was disclosed for a more limited purpose could compromise the privacy of the individuals involved.

Privacy, confidentiality, and security are closely related but distinct issues. As used here, *privacy* refers to an individual's interest in limiting the disclosure of personal information; *confidentiality* refers to controlling the release of information once it has been disclosed; and *security* refers to measures for controlling and protecting information and the systems through which it is accessed (National Research Council, 1997b). The fundamental concerns about unauthorized access to and use of health data are relevant for both paper records and electronic systems. The scope, power, and speed of electronic information systems magnify these concerns, but use of electronic information systems also offers new means for protecting data. Moreover, concerns about protecting the privacy, confidenti-

ality, and security of personally identifiable records apply not just to health data, but also to data collected for a variety of purposes, such as tax and census records.

In discussions of health data, medical records and administrative files that identify specific individuals are generally viewed as the most vulnerable to inappropriate disclosure, but other materials, such as vital statistics records and survey responses, also require adequate protection. Moreover, even supposedly anonymous or aggregated data, such as those published in vital statistics reports, must be handled appropriately because distinctive combinations of characteristics such as age, race, occupation, and diagnosis could permit the identification of individuals. Linking records across data sets can add valuable information, but poses added risks that privacy and confidentiality will be compromised. The proposed use of unique personal identifiers in health records discussed earlier would facilitate record linkage, but many observers oppose their adoption until more effective privacy protections are in place (e.g., Institute of Medicine, 1994; National Committee on Vital and Health Statistics, 1997b; National Research Council, 1997b).

Several federal laws and regulations provide privacy protection for data collected by federal agencies (see National Research Council, 1993), and states have adopted varying provisions regarding the privacy and confidentiality of publicly and privately held data related to health status (e.g., infectious disease reports) and health care (Institute of Medicine, 1994; Gostin et al., 1996). Many observers believe that federal legislation is necessary to ensure a uniform minimum level of protection for health data at the national, state, and local levels and in the public and private sectors (e.g., Institute of Medicine, 1994; National Committee on Vital and Health Statistics, 1997a). Others have advocated the adoption of more consistent policies governing the collection and use of data for statistical purposes by federal agencies (National Research Council, 1993).

Despite widespread support for strong protection of the privacy and confidentiality of most health data, considerations of personal or public safety may sometimes require controlled release of data related to matters such as contagious diseases or mental illness. Other public policy priorities may also preempt protections for health-related information. For example, states now have the option to bar permanently from the Temporary Assistance for Needy Families program persons convicted of drug-related felonies (U.S. Department of Health and Human Services, 1997d).

Over the past few years, Congress has considered but not acted on proposals to establish policies regarding the privacy and confidentiality of individually identifiable health data. Action may now be more likely because HIPAA calls for Congress to pass such legislation by August 1999, or if Congress does not act, for the Secretary of Health and Human Services to issue regulations for electronic administrative and financial transactions. DHHS has submitted recommendations to Congress for federal privacy standards (U.S. Department of Health and Human Services, 1997b), and discussion continues over broader federal action. In terms of performance measurement, the panel notes that the DHHS recom-

mendations include provisions for disclosure of information to public health agencies and state health data systems. The National Committee on Vital and Health Statistics (1997a) has strongly recommended passage of a health privacy law rather than reliance on departmental regulations that will govern only electronic transactions because the restricted scope of those regulations may make them impossible to administer appropriately.

Ensuring the physical security of health data, especially data in electronic form, is an essential adjunct to policies on privacy and confidentiality, but relates also to protecting data from intentional or inadvertent alteration. Although technological measures can increase data security, strong organizational policies and practices are also needed. Specific recommendations regarding health data have been made by a committee of the National Research Council (1997b) and the National Committee on Vital and Health Statistics (1997a). Among the technological steps that have been recommended to ensure the security of electronic health data are individual authentication of information system users, procedures to control user access to data, tracking and review of user transactions (i.e., use of "audit trails"), and protection of electronic communications and points of remote access to an information system. Among the recommended organizational practices are establishing formal security and confidentiality policies that include sanctions for violations, designating an information security officer, and providing training for staff and other users of an organization's information systems.

The policies and practices that emerge in response to these concerns are likely to have significant implications for health information systems and the use of health data for performance measurement. The panel urges careful consideration of all perspectives, recognizing that there are strong views and compelling interests on many sides. These issues lie beyond the scope of the panel's work, but policy makers and the public are urged to consider carefully the recommendations that others have made.

INVESTING IN HEALTH DATA AND DATA SYSTEMS

For performance measurement to be effective, good data and information systems, including a skilled staff with expertise to manage and use those systems, must be developed and maintained. A stable long-term investment must be made in the equipment and program activities needed to collect, manage, and use health data. To ensure that staff at the federal, state, and local levels are prepared to perform the tasks associated with performance measurement, a similar investment must be made in training and technical assistance. The DHHS strategic plan gives high priority to investments in electronic data systems and the training and technical assistance needed to apply new technologies at the state and local levels (U.S. Department of Health and Human Services, 1997a).

These investments are a responsibility that should be shared by all who expect to make use of the data and information systems involved, whether for perfor-

mance measurement or for other purposes. Support must include adequate direct funding, as well as commitments of staff time and access to computing and other technical resources. Because publicly funded health programs often face serious funding constraints, the panel emphasizes the importance of mobilizing the resources needed for data and information system development in ways that do not compromise funding for program services. At the same time, it is important to emphasize that only with good data and good program monitoring is it possible to assess whether program services are effective and being used appropriately.

Data and Information Systems

The panel urges DHHS to initiate a comprehensive review of the nation's current portfolio of health data activities to explore with states, communities, and the private sector opportunities for producing better data more efficiently. Although the panel has focused on performance measurement, this review should adopt a broader perspective that takes into account the variety of purposes served by health data and information systems.

Among the issues for exploration in this review are the investment in federal versus state and local data systems and opportunities for more efficient use of data system resources. For example, the DHHS strategic plan calls for additional investment in departmental surveys to generate state-level data, but the merits of using federal resources to strengthen state and local survey programs should also be considered. This latter approach might make it possible to consolidate reports to meet certain national data needs while producing data that respond to specific state and local needs as well.

A key starting point might be the BRFSS model. This collaboration between CDC and the states has resulted over time in a survey program in which all states and the District of Columbia participate. Each of the annual state-administered surveys uses a standard core questionnaire and can also include separately funded customized supplements that respond to specific state interests. These surveys also provide a framework that states can use to produce more detailed substate data. The BRFSS has been cited as a key source of state-level data for measures related to Healthy People (Kim and Keppel, 1997). It is also an essential resource for the state-level data on health status and health risk factors that will be needed for many performance measures, including a number of the measures proposed by this panel in its first report (National Research Council, 1997a). The panel is concerned that the BRFSS program has not received a strong commitment at the federal level for continuing support consistent with its importance as an information resource. The federal and state funding arrangements vary from state to state, but overall, direct federal funding has generally supported about half of the modest annual cost of the survey program. For the 1996-1997 grant cycle, this direct federal support amounted to about \$3.5 million of the combined federal-state funding of about \$7 million (D. Nelson, Centers for Disease Control and

Prevention, personal communication, September 1998). For the 1997-1998 grant cycle, however, direct CDC support was reduced by about 25 percent to \$2.7 million. A funding loss of this magnitude is a serious concern. For the 1998-1999 round of grants, CDC funding for the BRFSS increased to \$3.9 million, but even with this increase, the funding level allows for an average grant of only about \$76,500 per state.

Efforts to identify opportunities for making more efficient use of existing data resources could include assessing the usefulness of the data currently being collected, exploring opportunities to build new data collection capabilities within existing systems, and identifying ways to remove obstacles that may hinder more efficient operation or the sharing of data across systems. Given changes in program priorities and a more outcome-oriented approach to monitoring program operations (as reflected in the PPG proposal), some current data collection programs may no longer be appropriate or may require redirection. If out-of-date activities were identified, the resources used to support them could be shifted to more useful data activities.

Opportunities may also exist to expand data collection within an existing framework, which would tend to be less costly than establishing a new free-standing activity. For example, the National Center for Health Statistics has proposed the State and Local Area Integrated Telephone Survey (SLAITS), which would take advantage of the National Immunization Survey sampling frame. For the latter survey, a large number of ineligible households must be contacted in the course of identifying those with children of an appropriate age. With SLAITS, these contacts with ineligible households could be transformed into contacts with households eligible for alternative surveys.

Also of concern to the panel are data system inefficiencies that may exist because of constraints on the use of categorical funding or demands for specialized data systems and reporting. For example, states have found it difficult to integrate some federally developed reporting systems into existing state information systems. In a recent audit, the Illinois Department of Health found eight separate information systems for HIV/AIDS, each of which required independent data entry (J. Lumpkin, Illinois Department of Health, personal communication, August 1998). Specialized turnkey or proprietary systems that are customized for a single program area can be difficult to link to other information systems or adapt for other, related applications. Moreover, because such systems must be used in operational settings that vary across communities and states, a single version is unlikely to be suitable for every setting. Information science advises designing information systems to support service delivery rather than adapting service delivery to the information systems.

Problems of redundancy and incompatibility can be traced to all levels of government. If such problems can be identified, efforts can be made to overcome them, although they may not be easy to eliminate. At the federal level, CDC and the Health Resources and Services Administration (HRSA) recently took steps in

this direction by endorsing the use of their categorical grant funds to support the development of integrated health information systems, noting that integration will benefit categorical programs and serve cross-cutting information needs (Broome and Fox, 1998). The panel encourages other federal agencies, as well as state and local health agencies, to explore similar policies. Any formal legislative and regulatory restrictions that constrain the use of program funds in support of more integrated health information systems should be reviewed to determine whether they can be revised or removed.

Technology

Rapid advances in information technology are presenting new opportunities to collect, manage, analyze, and disseminate data for performance measurement and other purposes. To take full advantage of those opportunities, however, federal, state, and local governments must invest in more sophisticated computers, software, and communications capabilities. To optimize their investment, they should look for efficient approaches to system design and operation. For example, a modular object-oriented approach to programming facilitates the transfer of software development efforts from one application to another. This reuse of software can dramatically reduce the time and cost of system development. The creation of a national repository of software objects that perform common core functions might be one means of facilitating the development of state or local information systems and leveraging the funds currently available for system development.

The panel has not attempted to estimate the level of investment that would be appropriate, but notes that in the private sector, the health care industry spent an estimated \$10-\$15 billion on information technology in 1996 (Munro, 1996). Further growth in the level of effort is expected as health care organizations implement CPRs; upgrade administrative and billing systems; install networks for sharing information with affiliated entities; and use public networks, such as the Internet, to distribute health-related information and provide access to clinical databases in remote areas. The scale of private-sector investment signals broad recognition of the importance of supporting information systems.

If the public sector is not to be left behind, it, too, must make a significant investment in information systems. Estimates of the current spending on state and local health data systems are not readily available,² but a reference point

²The Association of State and Territorial Health Officials, the National Association of County and City Health Officials, the National Association of Local Boards of Health, and the Public Health Foundation are collaborating in a federally funded project aimed at developing a methodology for measuring state and local public health expenditures in support of the essential public health functions (see Chapter 3 for a list of these functions). The results of this project are expected to lead to better information about public health spending, but may not directly address investment in information systems. Descriptions of the project can be found at <www.naccho.org/projects/expend.html> and <www.phf.org/policy.htm#State/Local>.

might be sought in state spending on environmental data systems. The Environmental Protection Agency (EPA) (1998) reports that states engaged in reforming their environmental reporting processes are spending \$3-\$10 million per year for data system improvements and operations. EPA is making demonstration grants of up to \$500,000 to support these efforts. In DHHS, the Maternal and Child Health Bureau (1998) is funding a systems development initiative that offers states grants of \$100,000 that can be used to support information system activities, especially those related to performance measures for the Maternal and Child Health Services Block Grant. In the future, federal, state, and local health data systems may benefit from the savings expected to result from the administrative simplification required by HIPAA. Strategies to realize this potential and reinvest some portion of the savings in data systems and performance measurement for publicly funded health programs should be explored.

Training and Technical Assistance

The adoption of performance-based systems of accountability for publicly funded health programs will require staff who oversee and operate these programs to apply skills in planning and assessment that may be unfamiliar to them. Federal agencies have found the development of specific performance goals and the definition of related outcome measures to be among the most difficult challenges posed by GPRA (U.S. General Accounting Office, 1997). The use of performance measurement also draws further attention to the need for expertise in data analysis and in the design and operation of data collection and data management systems. Greater access to data and to more powerful computers and software makes it easier to perform more complex analyses, but also increases the importance of ensuring that users have sufficient skills and expertise to use these technologies appropriately. Moreover, in commenting on requirements for successfully implementing GPRA, the U.S. General Accounting Office (1996) has observed that staff will need skills in strategic planning, performance measurement, and use of performance information in decision making and that agencies should view training to develop these skills as a worthwhile investment. Support for training and technical assistance is essential to ensure that the necessary skills and expertise are available.

In theory, health departments and other health-related agencies might add new staff to obtain the expertise needed to support performance measurement and related activities. In practice, however, most states and communities have limited resources for hiring additional program and data system staff, and some may face other pressures to maintain or reduce staff size. Furthermore, the relatively low salaries traditionally offered by state and local health agencies can make it difficult to attract and retain highly trained staff. These limitations may be especially acute in technical areas. The rapid growth of the information technology industry, as well as the need to address the year 2000 problem for huge numbers

of computer systems, has placed a high premium on information technology skills. Given these constraints, access to training programs that can enhance the skills of existing staff and to technical assistance that draws on the expertise of others becomes especially important.

For staff to obtain the needed training, suitable materials and programs are required, as well as time and funds to support the staff members' participation. Training opportunities may take many forms, including formal academic programs (e.g., graduate programs in schools of public health) and specialized courses and training sessions offered by federal agencies (e.g., the CDC Public Health Training Network), academic institutions, or others in the private sector. Funding for scholarships and dissertation grants could assist staff in obtaining advanced academic training. Support for other training opportunities is also needed. Teleconferencing, self-guided instruction, and other forms of distance-based learning can bring a variety of training to large audiences and can compensate in part for constraints on funding for travel to attend courses and conferences. However, supplementing distance-based training with attendance at off-site programs may give staff valuable opportunities to learn through direct interaction with colleagues from other states or communities. The panel was informed that even though CDC provides funds specifically to allow staff from each state to attend the annual BRFSS conference, these funds generally cover participation by the data managers who oversee the collection and maintenance of state BRFSS data sets, but are not adequate to support the attendance of most users of BRFSS data.³

Technical assistance can make a large reservoir of expertise available to meet diverse needs. The assistance can take many forms, including publications, information clearinghouses, conferences, and consultations with experts. In a recent activity of particular relevance to the interests of this panel, CDC and HRSA worked with the Association of State and Territorial Health Officials and the National Association of County and City Health Officials to develop an "investment guide" to assist states in planning and developing integrated health information systems (Centers for Disease Control and Prevention and Health Resources and Services Administration, 1998).

A review of technical assistance activities in DHHS led to the conclusion that these activities could be enhanced by greater coordination and evaluation of the effectiveness of current forms of assistance (U.S. Department of Health and Human Services, 1997f). It was suggested to the panel that in the area of epidemiologic analysis, for example, states could benefit from greater access to more senior CDC epidemiologists to supplement programs that currently rely primarily

³This information was reported to the panel in the background paper "Improving Federal-State Data Collection to Monitor Program Performance Measures," which was prepared by the Science and Epidemiology Committee of the Association of State and Territorial Chronic Disease Program Directors and the Council of State and Territorial Epidemiologists.

on newly trained epidemiologists.⁴ Because of their national perspective and their influential role as funders of many health programs, federal agencies are well placed to serve as a focal point for technical assistance. User groups that draw participants from local, state, and federal health agencies could open other channels for obtaining technical assistance and learning about a broader range of health data issues. States and communities might also look to academic institutions and others in the private sector, particularly in a rapidly evolving area such as information technology. Foundations or other nonprofit groups might be able to serve as intermediaries in sponsoring such public-private collaborations.

TAKING A COLLABORATIVE APPROACH TO THE DEVELOPMENT OF HEALTH DATA AND INFORMATION SYSTEMS

The panel's deliberations regarding performance measurement have led to the conclusion that much greater collaboration and coordination are an essential foundation for further development of the nation's health data and data systems. It appears that by adopting a broadly based approach to health data needs and resources, it will be possible to make more effective use of available data and information systems for performance measurement, as well as for other purposes, including monitoring health status in the population, managing health programs, and informing policy makers and the public. For publicly funded health programs, it is essential that information needs at the federal, state, and local levels all be taken into account. The DHHS strategic plan recognizes the need for accurate and timely data at all these levels for assessing changes in health status and managing health programs (U.S. Department of Health and Human Services, 1997a).

States are responding to these concerns with initiatives aimed at strengthening their health data infrastructure by improving data quality; developing standards for data definitions, information system configurations, and electronic transmission of data; and linking data systems (see the earlier discussion in this chapter).⁵ The panel is encouraged to see states taking these steps and believes there is additional value in promoting a national approach to these matters. State-specific solutions may limit the comparability of data across states, and states may miss opportunities to collaborate or to adopt successful strategies developed elsewhere. Likewise, the panel applauds the advances that HIPAA is expected to bring to standards for electronic health care transactions, but also urges support for efforts that will encourage the development of standards for an even broader range of health data elements, such as those likely to be used in performance measures for a variety of publicly funded health programs.

⁴This suggestion was also made to the panel in the background paper "Improving Federal-State Data Collection to Monitor Program Performance Measures."

⁵A summary of state efforts to integrate health information was compiled by DHHS and The Lewin Group. Information on activities in each state can be found at <<http://aspe.os.dhhs.gov/statereg/>>.

Meeting the Needs of Many Data Users

Many federal health data systems have been designed to provide national-level data for an overall assessment of health status to help guide the planning and implementation of national health policies and programs. At the state and local levels—the “front line” for service delivery—the perspective is somewhat different. Detailed local data are needed to guide planning and program operations, and they have more immediate value than national estimates. Even summary state-level data may lack sufficient detail to be useful for understanding health needs and program outcomes at the local level. For example, data for Illinois as a whole are not likely to provide a satisfactory picture of health status and program activities in either Chicago or a rural county in southern Illinois. Developing a more efficient and effective approach to information systems used to support performance measurement may depend on finding a way to accommodate differing perspectives on several issues.

One concern is the tension between the program-specific perspective that is often the basis for funding and oversight of publicly funded health programs and a more functional perspective on the operation of data systems that focuses on the commonalities among the data collection and management tasks to be performed for many program areas. Categorical grant programs help ensure that funds are directed to specific needs, but they may hinder both a broad view of health and the efficient organization of data systems at the state and local levels.

At the federal level, the programmatic perspective often dominates. The various categorical funding programs often have specialized reporting requirements, and some require the use of independent, customized systems to file those reports (e.g., for HIV/AIDS cases as noted earlier). In contrast, an approach that consolidates data collection systems across program areas can be beneficial at the state and local levels, where limited staff and operational funding can be used more efficiently if similar tasks can be combined. For example, a single ongoing survey such as a state’s Behavioral Risk Factor Survey can collect data on such topics as smoking habits, alcohol use, disabling conditions, and mammography use without requiring each program to operate a separate survey.

The specialized data systems developed to meet categorical program requirements tend to have a limited scope and may be costly to maintain. They may require duplication of data collection and management tasks, and if their reporting requirements are incompatible, they may preclude use of a single, more efficient data collection method at the state or local level. Cooperation across programs can provide an opportunity to combine resources from diverse program areas to support similar tasks in data collection and analysis. For example, a single system of notifiable disease surveillance could accommodate reports on AIDS cases and pesticide exposures, or a single ongoing telephone survey of adults could integrate questions about domestic violence and mammography use. For this approach to work, compromises may be needed to balance the interests of diverse program areas. If a single survey that addresses both domestic violence

and mammography use is to remain short enough to be practical, it may have to collect less detail on each topic than would be gathered by separate surveys.

State and local officials and health planners are also concerned about the flexibility and timeliness of data collection and reporting. They require access to current information about the specific populations they serve for effective program implementation and management. Often, data systems managed at the federal level have not been able to respond to these needs. For example, the National Health Interview Survey, National Health and Nutrition Examination Survey, and National Hospital Discharge Survey produce valuable national data, but are not designed to produce state- or local-level estimates. Moreover, state and local health departments and other health-related agencies have had little organized opportunity to participate in shaping the design and content of many federally operated data systems. Without this input, such systems are less likely to be relevant to state and local concerns, and opportunities to improve comparability or coordination across federal, state, and local data systems may be missed. Also, data managed at the national level have often been produced more slowly than is useful for state and local purposes. New computer and communications technologies are reducing the time needed to collect and process data and produce reports, but they may require expertise and equipment that are not yet available in some states and communities. To the extent that federal data systems will be relied upon to meet the need for state and local data for performance measurement, those systems will require the capacity and flexibility to respond in a timely way to state or local information needs. They will also have to ensure that data processing and reporting proceed as expeditiously as possible.

The need for timely data of state or local relevance should not, however, undermine the quality of the data in terms of validity, reliability, completeness, or accuracy. For example, new survey questions or modules must be validated, and survey staff must be trained to administer them. Concerns at the federal level about the quality and comparability of data produced by states have tended to encourage federal centralization of data systems rather than aggregation of state-level data. Although states acknowledge shortcomings in some areas, they are committed to producing high-quality data. Federal-state collaborations in areas such as vital records data and AIDS case reporting have achieved good quality and comparability in state-based data systems. These collaborations stand as examples for efforts that could be undertaken in other areas, such as enhancing the comparability of states' behavioral risk factor data.

Another source of tension is the burden associated with the reporting requirements for the federal block grants (e.g., the Preventive Health and Health Services Block Grant or the Substance Abuse Prevention and Treatment Block Grant) that provide a portion of the funds used to support state and local health programs. In the past, the reporting requirements associated with these grants have imposed a significant burden because some of the required information is not readily available and is often expensive or time-consuming to obtain. In

addition, the reporting requirements of different federal grants are not always consistent across program areas. Constraints on the use of grant funds have also tended to prevent the consolidation of funding to support the development of integrated data systems. The performance partnership concept represents an effort to reduce this burden by making the states partners in a negotiation with the granting agency that leads to the selection of some of the measures to be reported. Plans should be made to assess the impact of this approach on states' reporting burden.

Collaboration in the Design and Implementation of Data Systems

The panel has concluded that a more collaborative approach to the planning, design, and operation of health data systems would better serve the needs of all parties at the federal, state, and local levels. This conclusion is consistent with the views of the Council of State and Territorial Epidemiologists (1997a,b), as reflected in that organization's recommendations in support of a National Public Health Surveillance System and for enhanced usefulness of state and local data collection by the National Center for Health Statistics. Those recommendations included improving access to surveillance data through better coordination of data systems, and planning surveillance and other data collection activities at the state and local levels in a standardized but collaborative fashion that includes local, state, and federal partners from relevant organizations.

The panel's position is also consistent with the follow-up steps proposed as a result of the 1997 review of progress toward the Healthy People 2000 objectives on surveillance and data systems (U.S. Department of Health and Human Services, 1997e). Those proposals included involving state and local governments at every stage of national data collection, analysis, and dissemination; providing easier access to national data sets, including additional geocoding to facilitate subnational analyses; improving coordination of data resources within DHHS and between census and health program data; and giving greater attention to state and local priorities in the development of health objectives for Healthy People 2010.

Collaborative efforts are complicated by the multiplicity of stakeholders across the federal, state, and local levels. No single voice at any of these levels can speak to all of the issues that need to be addressed, and no established framework is currently available for selecting representatives and involving them in deliberations about data system issues (e.g., survey design, question selection). At the federal level, DHHS has a critical leadership role to play in these activities, but it must function as a partner with other stakeholders. Mechanisms are needed for designating recognized representatives of key stakeholder groups and for supporting their participation in formal and informal efforts to improve coordination and collaboration. Currently, opportunities for state officials to meet with their federal counterparts may be lost because funding constraints prevent out-of-state

travel. Similarly, states must work in partnership with community-level stakeholders, as well as with relevant federal and private-sector groups, to ensure that community information needs are addressed. For state and local government, the stakeholders include both staff with policy and programmatic responsibilities who use health data and staff with technical expertise in data collection and analysis who produce and manage health data.

Collaboration must be pursued not only in an intergovernmental framework, but also intragovernmentally. Better coordination among federal agencies, both within DHHS and between DHHS and other departments, could contribute to more effective use of available data and data collection systems and help reduce duplication of reporting effort for states and communities. Similarly, greater collaboration among states and communities increases the likelihood that they will be able to learn from each other and develop comparable measures, definitions, and data collection methods for monitoring health programs.

CONCLUSIONS

Health-related data are needed for the formulation of health policies and for the optimal targeting of resources to address priority health issues. Recent interest in performance measurement and performance-based accountability has brought renewed and broader attention to many long-standing concerns about these data and the data systems through which they are produced and used. The panel is convinced that this interest could and should be translated into the sustained commitment of time and resources needed to develop a more comprehensive and coherent approach to health data and health data systems that would build effectively on existing data resources and be capable of meeting health information needs at the federal, state, and local levels. The panel has focused primarily on the public-sector perspective, but recognizes that there are closely related private-sector interests and developments that must not be overlooked.

Attention must be given both to operational concerns and to policy issues. On the operational side, one of the most fundamental requirements must be ensuring that good-quality data are available and used in appropriate analyses. To make health data more useful in a broader context, greater consistency and comparability are needed. Key to achieving this objective will be the variety of activities under way to establish standards for the methods used to collect the data; the content and format of data files; the formats for exchanging data electronically; the protection of data privacy, confidentiality, and security; and the measures used to assess performance. Advances in computer technology and electronic data transmission could speed the collection and analysis of data and facilitate access to a broader range of health-related data for many more users.

The fundamental need is for a collaborative partnership across the local, state, and federal levels as a basis for strengthening and better coordinating the health data and information systems needed to support performance measurement.

Mechanisms must allow stakeholders to participate in an ongoing process that encourages them to contribute to policy determinations about what information is to be collected and how it is to be used. Because health issues affect everyone and are addressed in a variety of ways, the panel supports a national approach that recognizes a broad range of interests.

DHHS has an important leadership role to play in furthering these efforts, but all of the participants must share responsibility for ensuring that health data and data systems receive the support they need to operate efficiently and effectively. An investment must be made in the data collection programs and information technology that are at the core of these information systems and in the necessary training and technical assistance for the people who produce and use health data.

REFERENCES

Anderson, J.E., D.E. Nelson, and R.W. Wilson
1998 Telephone coverage and measurement of health risk indicators: Data from the National Health Interview Survey. *American Journal of Public Health* 88:1392-1395.

Arday, D.R., S.L. Tomar, D.E. Nelson, R.K. Merritt, M.W. Schooley, and P. Mowery
1997 State smoking prevalence estimates: A comparison of the behavioral risk factor surveillance system and current population surveys. *American Journal of Public Health* 87:1665-1669.

Bailar, J.C., and F. Mosteller, eds.
1992 *Medical Uses of Statistics*, 2nd ed. Boston: NEJM Books.

Broome, C.V., and C.E. Fox
1998 CDC/HRSA Grant Funding Flexibility for Integrated Health Information Systems. Grant funding transmittal letter. April 1, 1998. U.S. Department of Health and Human Services. <http://www.hrsa.dhhs.gov/policy.htm> (also at <http://www.cdc.gov/funds/policy.htm>) (April 21, 1998).

Centers for Disease Control and Prevention
1997 Case definitions for infectious conditions under public health surveillance. *MMWR* 46(RR-10).

Centers for Disease Control and Prevention and Health Resources and Services Administration
1998 Integrated Health Information Systems Investment Analysis Guide. <http://www.hrsa.dhhs.gov/investment.htm#iv> (also at <http://www.cdc.gov/funds/invest7.htm>) (April 21, 1998).

Clarke, K.C., S.L. McLafferty, and B.J. Tempalski
1996 On epidemiology and geographic information systems: A review and discussion of future directions. *Emerging Infectious Diseases* 2(2):85-92.

Council of State and Territorial Epidemiologists
1997a Implementation of National Public Health Surveillance System. Position statement #EC-1. Adopted June 19, 1997. <http://www.cste.org/page9.html> (September 28, 1998).

1997b NCHS-State Data Coordination. Position statement #CD-2. Adopted June 19, 1997. <http://www.cste.org/page58.html> (September 28, 1998).

Environmental Protection Agency
1998 One Stop Program Strategy and Grant Award Criteria. <http://www.epa.gov/reinvent/onestop/strategy.htm> (April 21, 1998).

Foundation for Accountability

- 1998 About FACCT. <http://www.facct.org/about.html> (April 15, 1998).

Gordis, L.

- 1996 *Epidemiology*. Philadelphia: Saunders.

Gostin, L., Z. Lazzarini, V.S. Neslund, and M.T. Osterholm

- 1996 The public health information infrastructure: A national review of the law on health information privacy. *JAMA* 275:1921-1927.

Hoaglin, D.C., R.J. Light, B. McPeck, F. Mosteller, and M.A. Stoto

- 1982 *Data for Decisions: Information Strategies for Policymakers*. Cambridge, Mass.: Abt Books.

Iezzoni, L.I., ed.

- 1994 *Risk Adjustment for Measuring Health Care Outcomes*. Ann Arbor, Mich.: Health Administration Press.

Iezzoni, L.I.

- 1997a Assessing quality using administrative data. *Annals of Internal Medicine* 127:666-674.
 1997b The risks of risk adjustment. *JAMA* 278:1600-1607.

Institute of Medicine

- 1994 *Health Data in the Information Age: Use, Disclosure and Privacy*. M.S. Donaldson and K.N. Lohr, eds. Committee on Regional Health Data Networks. Washington, D.C.: National Academy Press.
 1997 *The Computer-Based Patient Record: An Essential Technology for Health Care*, Revised ed. R.S. Dick, E.B. Steen, and D.E. Detmer, eds. Committee on Improving the Patient Record. Washington, D.C.: National Academy Press.

Joint Center for Poverty Research

- 1998 *Administrative Data for Policy-Relevant Research: Assessment of Current Utility and Recommendations for Development*. V.J. Hotz, R. Goerge, J.D. Balzekas, and F. Margolin, eds. Report of the Advisory Panel on Research Uses of Administrative Data. Evanston, Ill.: Northwestern University/University of Chicago Joint Center for Poverty Research.

Joint Commission on Accreditation of Healthcare Organizations

- 1998 Performance Measurement. <http://www.jcaho.org/perfmeas/perfmsmt.htm> (July 24, 1998).

Kim, I., and K.G. Keppel

- 1997 *Priority Data Needs: Sources of National, State, and Local-Level Data and Data Collection Systems*. Healthy People 2000 Statistical Notes, No. 15. Hyattsville, Md.: National Center for Health Statistics.

Landon, B., L.I. Iezzoni, A.S. Ash, M. Shwartz, J. Daley, J.S. Hughes, and Y.D. Mackiernan

- 1996 Judging hospitals by severity-adjusted mortality rates: The case of CABG surgery. *Inquiry* 33:155-166.

Lasker, R.D., B.L. Humphreys, and W.R. Braithwaite

- 1995 Making a Powerful Connection: The Health of the Public and the National Information Infrastructure. Report of the U.S. Public Health Service Public Health Data Policy Coordinating Committee. Washington, D.C. <http://www.nlm.nih.gov/pubs/staffpubs/lo/makingpd.html> (August 11, 1998).

Luft, H.S., and P.S. Romano

- 1993 Chance, continuity, and change in hospital mortality rates. Coronary artery bypass graft patients in California hospitals, 1983 to 1989. *JAMA* 270:331-337.

Maternal and Child Health Bureau

- 1998 State Systems Development Initiative (SSDI) Grant Application Guidance for FY98. U.S. Department of Health and Human Services, Health Resources and Services Administration. <http://www.hhs.gov/80/hrsa/mchb/guidance.htm> (June 4, 1998).

McLellan, A.T., H. Kushner, D. Metzger, R. Peters, I. Smith, G. Grisson, H. Pettinati, and M. Argeriou
 1992 The fifth edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment* 9:199-213.

Mendelson, D.N., and E.M. Salinsky
 1997 Health information systems and the role of state government. *Health Affairs* 16(3):106-119.

Munro, N.
 1996 Infotech reshapes health care marketplace. *Washington Technology* 11(August 8):1-3.

National Association of State Mental Health Program Directors
 1998 State Mental Health Directors Adopt a Framework of Performance Indicators for Mental Health Systems. Press announcement. May 1998. Alexandria, Va.

National Association of State Mental Health Program Directors Research Institute
 1998 *Five State Feasibility Study on State Mental Health Agency Performance Measures*. Final Report. Alexandria, Va.: National Association of State Mental Health Program Directors Research Institute.

National Committee for Quality Assurance
 1996 NCQA Issues Final Technical Specifications for HEDIS 3.0. Press release. October 31, 1996. <http://www.ncqa.org/news/techrel2.htm> (April 7, 1998).

1997 *HEDIS 3.0/1998. Vol. 4, A Roadmap for Information Systems: Evolving Systems to Support Performance Measurement*. Washington, D.C.: National Committee for Quality Assurance.

National Committee on Vital and Health Statistics
 1997a Health Privacy and Confidentiality Recommendations. June 25, 1997. Washington, D.C. <http://aspe.os.dhhs.gov/ncvhs/privrecs.htm> (June 11, 1998).

1997b Letter to Secretary Shalala on Unique Health Identifiers. September 9, 1997. Washington, D.C. <http://aspe.os.dhhs.gov/ncvhs/uhid.htm> (June 11, 1998).

National Highway Traffic Safety Administration
 1996 *Why Data Linkage? The Importance of CODES*. Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration. Available at <http://www.nhtsa.dot.gov/people/ncsa/codes/whylink.html> (April 28, 1998).

National Library of Medicine
 1998 Fact Sheet: Unified Medical Language System. National Institutes of Health. February 23, 1998. <http://www.nlm.nih.gov/pubs/factsheets/umls.html> (July 10, 1998).

National Research Council
 1993 *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics*. G.T. Duncan, T.B. Jabine, and V.A. de Wolf, eds. Panel on Confidentiality and Data Access, Committee on National Statistics. Washington, D.C.: National Academy Press.

1997a *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health*. E.B. Perrin and J.J. Koshel, eds. Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics. Washington, D.C.: National Academy Press.

1997b *For the Record: Protecting Electronic Health Information*. Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure, Computer Science and Telecommunications Board. Washington, D.C.: National Academy Press.

1998 *Providing National Statistics on Health and Social Welfare Programs in an Era of Change*. Summary of a workshop. C.F. Citro, C.F. Manski, and J. Pepper, eds. Committee on National Statistics. Washington, D.C.: National Academy Press.

- Office of the Secretary, U.S. Department of Health and Human Services
 - 1998 National standard health care provider identifier. Proposed rule. May 7. *Federal Register* 63(88):25320-25357.
- Palmer, R.H.
 - 1997 Process-based measures of quality: The need for detailed clinical data in large health care databases. *Annals of Internal Medicine* 127:733-738.
- Powe, N., J. Weiner, B. Starfield, M. Stuart, A. Baker, and D. Steinwachs
 - 1996 The development and testing of a claims data-based approach for evaluating the care of patients with chronic illnesses. *Medical Care* 34:798-810.
- Public Health Foundation
 - 1998 *Measuring Health Objectives and Indicators: 1997 State and Local Capacity Survey*. Washington, D.C.: Public Health Foundation.
- Ries, L.A.G., C.L. Kosary, B.F. Hankey, A. Harras, and B.K. Edwards, eds.
 - 1997 *SEER Cancer Statistics Review, 1973-1994*. NIH Pub. No. 97-2789. Bethesda, Md.: National Cancer Institute.
- Rothman, K.J.
 - 1986 Stratified analysis. Pp. 176-236 in *Modern Epidemiology*. Boston: Little, Brown.
- Rubin, H., M.W. Jenckes, M. Stuart, and M. Wickham
 - 1994 Maryland Medicaid Recipients' Ratings of Prenatal and Pediatric Care by HMOs and Fee-for-Service Providers. Paper presented at 122nd meeting of the American Public Health Association (abstract, p. 324). Washington, D.C.
- Seitz, F., and B. Jonas
 - 1998 *Operational Definitions for Year 2000 Objectives: Priority Area 6, Mental Health and Mental Disorders*. Healthy People 2000 Statistical Notes, No. 16. Hyattsville, Md.: National Center for Health Statistics.
- Starfield, B., N. Powe, J. Weiner, M. Stuart, D. Steinwachs, S. Scholle, and A. Gerstenberger
 - 1994 Costs vs. quality in different types of primary care settings. *JAMA* 272(24):1903-1908.
- Starr, P.
 - 1997 Smart technology, stunted policy: Developing health information networks. *Health Affairs* 16(3):91-105.
- Starr, P., and S. Starr
 - 1995 Reinventing vital statistics: The impact of changes in information technology, welfare policy, and health care. *Public Health Reports* 110:534-544.
- Strouse, R., J. Hall, B.L. Carlson, and J. Cheng
 - 1997 Impact of the Nontelephone Sample on Family Health Insurance Survey Estimates. Report submitted by Mathematica Policy Research, Inc., to the Robert Wood Johnson Foundation. March 3, 1997. Princeton, N.J.
- Stuart, M.
 - 1994 Redefining boundaries in the financing and care of diabetes: The Maryland experience. *Milbank Quarterly* 72(4):679-694.
 - 1995 Public health issues in the development of centralized health care databases. *Journal of Public Health Management and Practice* 1(2):31-38.
- Stuart, M., D. Steinwachs, J. Harlow, and M. Fox
 - 1990 Ambulatory practice variation in Maryland: Implications for Medicaid cost management. *Health Care Financing Review* (Annual) December:57-67.
- Svikis, D., M. McCaul, T. Feng, M. Stuart, M. Fox, and E. Stokes
 - 1998 Drug dependence during pregnancy: Effect of an on-site support group. *Journal of Reproductive Medicine* 43:799-805.
- Thacker, S.B., and D.F. Stroup
 - 1994 Future directions for comprehensive public health surveillance and health information systems in the United States. *American Journal of Epidemiology* 140:383-397.

Turner, C.F., L. Ku, S.M. Rogers, L.D. Lindberg, J.H. Pleck, and F.L. Sonenstein
1998 Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science* 280(5365):867-873.

U.S. Department of Health and Human Services
1991 *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. DHHS Pub. No. (PHS) 91-50212. Washington, D.C.: Office of the Assistant Secretary for Health.

1997a 1997 Strategic Plan. September 30, 1997. Washington, D.C.

1997b Confidentiality of Individually-Identifiable Health Information. Recommendations of the Secretary of Health and Human Services, pursuant to Section 264 of the Health Insurance Portability and Accountability Act of 1996. September 11, 1997. Washington, D.C.

1997c Health Insurance Portability and Accountability Act of 1996 Administrative Simplification. Fact sheet. February 13, 1997. <http://www.hhs.gov/news/press/1997pres/970213.html> (June 29, 1998).

1997d Major Provisions of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (P.L. 104-193). April 28, 1997. <http://www.acf.dhhs.gov/news/welfare/aspesum.htm> (July 8, 1998).

1997e Progress review: Surveillance and data systems. *Prevention Report* 12(1).

1997f Technical Assistance in the U.S. Department of Health and Human Services. Report of the Technical Assistance and Training Liaison Work Group. July 1997. Washington, D.C. <http://aspe.os.dhhs.gov/progsys/ta/hhs-ta.htm> (April 13, 1998).

1998a *Leading Indicators for Healthy People 2000*. Report from the HHS Working Group on Sentinel Objectives. Washington, D.C.: U.S. Department of Health and Human Services.

1998b Registry of State-Level Efforts to Integrate Health Information. <http://aspe.os.dhhs.gov/statreg/index.htm> (February 9, 1998).

U.S. General Accounting Office
1996 *Executive Guide: Effectively Implementing the Government Performance and Results Act*. GAO/GGD-96-118. Washington, D.C.: U.S. Government Printing Office.

1997 *Managing for Results: Analytic Challenges in Measuring Performance*. GAO/HEHS/GGD-97-138. Washington, D.C.: U.S. Government Printing Office.

U.S. Public Health Service
1979 *Healthy People: Surgeon General's Report on Health Promotion and Disease Prevention*. DHEW (PHS) Pub. No. 79-55071. Washington, D.C.: U.S. Department of Health, Education, and Welfare.

5

Strategies for Supporting Performance Measurement Through a National Health Information Network

As the previous chapters have shown, further advances in performance measurement for publicly funded health programs will require thoughtful and continuing attention to a varied set of policy, programmatic, and data and information system issues. The current focus on performance-based accountability has helped highlight the limitations of existing health-related data and data sources. Despite their diversity and breadth, the data suitable for performance measurement are limited. In fact, the U.S. Department of Health and Human Services (DHHS) (1997:10) has acknowledged that lack of suitable data is a “critical limiting factor” in developing departmental performance objectives as required under the Government Performance and Results Act (GPRA). The need for stronger and more coherent data systems for public health surveillance and personal health services is a concern at the federal, state, and local levels and in the private sector as well (e.g., Thacker and Stroup, 1994; Gold, 1995; Lasker et al., 1995; Mendelson and Salinsky, 1997; National Committee for Quality Assurance, 1997; Starr, 1997; U.S. Department of Health and Human Services, 1997; Public Health Foundation, 1998).

This final chapter reviews the panel’s essential conclusions regarding performance partnership agreements, outlines its vision for a national health information network to facilitate performance measurement for publicly funded health programs, and recommends steps that can and should be taken to realize that vision. Although much in these conclusions and recommendations is applicable in principle to the broadest spectrum of health services in the public and private sectors, the panel has anchored its positions in the context of its discussions of

publicly funded health programs, including those in mental health, substance abuse, and various areas of public health.

CONCLUSIONS REGARDING PERFORMANCE PARTNERSHIP AGREEMENTS

Examination of performance measurement in the context of the proposal for Performance Partnership Grants (PPGs) for health programs has led the panel to a set of general principles that it believes should guide further performance partnership efforts. First, the panel concludes that those who are affected by decisions resulting from the application of performance measures must share fully in the creation and selection of those measures, and, where possible, the measurement process. Representatives from federal, state, and local health and health-related agencies, and often segments of the general public, various private-sector interests, and health care provider groups, should all have a significant voice in the development, selection, assessment, interpretation, and use of performance measures.

Second, the panel concludes that state- and local-level data are essential to the achievement of federal goals for performance measurement and that support of data collection mechanisms to produce those data is important to the success of the system. Most national surveys, for example, are not able to produce state-specific estimates that can be used to compare state-level performance. The development of performance measures and the data systems on which they depend should be approached with a broad national perspective that considers the interdependencies across governmental levels.

Third, the panel believes that performance measurement information, resources, and processes should be organized so that states and communities can avoid unnecessary duplication of effort in developing new or enhanced data systems by using, to the extent possible, existing data systems that already serve other purposes and by working together to learn from each other. Collaborative efforts within and among states can reduce the unnecessary inconsistencies and incompatibilities that tend to arise in independently developed monitoring programs or data systems. The challenge is to foster this collaboration and cooperation while preserving the flexibility needed by individual states and communities to accommodate diversity in their programs and goals.

Finally, the panel stresses that performance measurement should focus on the overall goals of an activity, not seek to measure primarily the impact of a particular source of funding for that activity. Generally speaking, program goals are best represented by outcome measures, and most health outcomes are influenced by many more factors than those that might be linked to a single funding source, making inferences concerning individual funding sources problematic.

A NATIONAL INFORMATION NETWORK FOR HEALTH-RELATED DATA

Having considered both the general features of performance measurement and the specific context of publicly funded health-related programs, the panel has concluded that over the long term, performance measurement for such programs will be made easier and more effective by the development of a broadly based national health information network that can promote a collaborative and coordinated approach across the local, state, and federal levels and can help in meeting diverse needs for a variety of health-related information, including performance data. This approach does not require the creation of an entirely new data system or a federally managed system. Instead, the panel envisions building on existing data systems operated by agencies at all levels of government, as well as looking to data systems in the private sector, to produce the information required for performance measurement.

The challenge is to develop a reasonably efficient and effective network from the current diversity of data sources and information needs. A collaborative approach, consistent with the partnership element of the PPG proposal, will be essential for success. Collaboration is needed to accommodate the missions of both population-based and personal health services and to facilitate harmonization of parallel activities that are conducted independently by individual states or communities (e.g., vital records systems, the Behavioral Risk Factor Surveillance System [BRFSS]). Among the most important tenets guiding the development of the envisioned information network should be the requirements that it meet real managerial and accountability needs; that it reflect important interdependencies and relationships across governmental and programmatic lines; and that it recognize that multiple, specialized data systems may no longer be affordable or consistent with other critical priorities.

In the current technological environment, an information network that can facilitate the transmission and aggregation of data from multiple sources without requiring the use of specific equipment or software is more feasible than ever before. Cooperation and collaboration are required, however, to establish agreement on such matters as electronic interfaces and data definitions. Effective use of technology also requires investments both in the development of staff expertise and in hardware and software. (See the discussion of this issue in Chapter 4.) Such investments have, however, been beyond the reach of many state and local health agencies.

The panel recognizes that developing a truly comprehensive information network of national scope is a massive, potentially overwhelming endeavor. The substantial overlap that the panel found in potential performance measures for the substantive areas it considered in its first report clearly argues against a narrow, disease-specific approach to data collection and analysis. However, the larger issues involved in developing a national information system of relevance to a

broad range of health and health care concerns will be much more difficult and complicated, both technically and politically, than the further development of disease-specific systems. This much broader undertaking will require the involvement of many DHHS agencies (e.g., Centers for Disease Control and Prevention [CDC], Health Resources and Services Administration [HRSA], Substance Abuse and Mental Health Services Administration, Food and Drug Administration, Health Care Financing Administration, National Institutes of Health), other federal agencies (e.g., Department of Agriculture, Environmental Protection Agency, Department of Education, Department of Housing and Urban Development, Department of Justice, Department of Transportation), and a similar array of state and local agencies, plus the private sector.

Establishing and maintaining the network envisioned by the panel will require a strong and continuing commitment by health agencies at the local, state, and federal levels. Even with such a commitment, this undertaking will almost certainly require an incremental approach. Many states do not yet have the capability to produce the information that would be required in any evaluation framework, and other states with more advanced data systems will have to make some changes to achieve compatibility with new standards that will evolve for a broad national network. Initial efforts might focus on enhancing existing information systems that serve either broad surveillance purposes or specific programmatic areas (e.g., children's health, substance abuse, mental health, chronic disease) or perhaps on conducting state or local demonstration projects to test the coordination of data systems across programmatic lines. Many of these public-sector programs must also rely on currently inadequate clinical data systems. Various observers (e.g., Institute of Medicine, 1997a; National Committee for Quality Assurance, 1997; Starr, 1997) have addressed the need for a commitment by health care organizations to make incremental improvements toward building a comprehensive information framework that can increase the health care industry's capacity to measure and improve performance. Successful development of information systems for performance measurement in the public or private sectors will require both time and financial investment from a variety of sources, as well as commitment and persistence.

ESSENTIAL FEATURES OF A HEALTH INFORMATION NETWORK TO SERVE DIFFERENT LEVELS OF DECISION MAKING

Given the complexity and scope of the task of developing the envisioned health information network, the panel believes that a clear vision of the intended long-term goal is required if the effort is to stay on course amid the vast amount of detail, variation, and difficult choices involved. Thus, the panel's vision for a cohesive, national health information network includes the following features.

National Collaboration The network should be developed through collabora-

ration among local, state, federal, and private efforts. A collaborative approach, in contrast to a more hierarchical or centralized model, requires the recognition of all participants as partners who can contribute to the success of the information network and the performance measurement activities it is intended to support, and whose information needs and program priorities must be taken into consideration. Participants should expect both to gain benefits from access to the information network and to assume responsibility for contributing to its effective operation, including ensuring the availability of adequate resources (e.g., funding, staff, data, information technology and expertise, commitment of policy makers).

With multiple participants at each level, a collaborative approach is necessary to promote efficient and effective data collection and use, as well as agreement on appropriate performance measures. The aim is to achieve a network that is national in scope without being specifically a federal enterprise. The voluntary state-based National Vital Statistics System (National Center for Health Statistics, 1998) serves as an example of a mature nationally collaborative network.

Linkage but not Consolidation The information network should be based on compatible structure and architecture to promote linkage of comparable data and sharing of information within and across the local, state, and national levels in a manner that is consistent with appropriate protections for personal privacy and the confidentiality and security of health-related data. The concept of a comprehensive network implies neither the desirability nor the need for a single national health database or information system.

Ability to Meet the Needs of Varied Users The network should, in the aggregate, support multiple purposes, including monitoring for performance-based accountability of population-based and personal health services, operation of health programs, and delivery and management of clinical care (including payment for services). To do so, it must encompass the full range of health-related services, from population-based and clinical preventive services through treatment and remediation. The network should be able to serve the information needs of managers, planners, health care providers, evaluators, policy makers, and the public at the national, state, and local levels. For example, data systems operated by states and communities can provide essential geographic detail and flexibility in data collection and analysis that are often not available from federal data systems, whereas federal data systems provide the broad national information that guides federal policy making and can serve as a reference point for assessing progress by individual states and communities.

For performance measurement specifically, the information network must be able to accommodate the differing data and information needs at various oversight levels. Some measures and data are important primarily or exclusively at the operating level, while other, often more limited information is appropriate for a more distant accountability audience. Plans for performance measurement should distinguish among these information needs and match the selection of measures accordingly.

Standardization of Data and Measures Critical to such a comprehensive network is the establishment of clear and common definitions of data elements, measures, and coding systems, and standard approaches to data collection so that information can be aggregated across multiple populations and regions and so that comparisons, where appropriate, can be made among populations and geographic areas.

The BRFSS, for example, illustrates the development of a standard survey questionnaire that is used independently in each state. Currently, this survey provides the primary means by which state-specific estimates of key behavioral risks are generated, but variations in survey methods among the states can limit the comparability of the data collected. To meet the full range of performance measurement requirements, new or modified data collection instruments with demonstrated reliability and validity or more standardized protocols may be needed to maximize the value of the information collected.

Appropriate Performance Measures For performance measurement, an information network should focus on measures of health outcomes and risk factors and on measures of processes and capacity that are widely recognized as linked to important health outcomes. For many program areas, the measures should be selected to represent multiple perspectives, specifically including that of the consumer (e.g., satisfaction with access to services). The range of measures must also be applicable to varied types of services (e.g., individual and population-based, inpatient and outpatient), to different age groups (e.g., youth, adult, and older adult), and to people of differing ethnic and cultural backgrounds.

Efficient and Effective Use of Resources The envisioned national health information network will require sufficient resources to operate effectively and support a wide range of users, but it should be organized efficiently to minimize its consumption of resources and any competition for resources with health programs and services. One consideration is efficient data collection. Certain types of data that can be used for performance measurement can be collected using various methods, such as consumer surveys, medical record reviews, or compilation of data from administrative records. Likewise, data might be collected at varying intervals (e.g., continuously, annually, every 2 years). In setting performance measurement requirements, careful judgments should be made about the intensity of the information required in terms of frequency of data collection, level of detail, and completeness of coverage (e.g., sample or census data).

Coordination and integration of information systems will often prove more efficient than the development and operation of program-specific systems. An information network should also take advantage of appropriate information available from sources that are not primarily health related. For example, mental health programs may want information from corrections department databases on the number of people with serious mental illness in youth and adult corrections systems. Training and technical assistance must be an integral part of the frame-

work for an information network so that staff at all levels have the skills to manage data systems effectively and use the information that they produce.

Adaptability to Change The national health information network should be established on the premise that it is a developing tool that must be able to adapt easily and in a timely fashion to changes and enhancement as information sources grow, knowledge expands, or requirements change. One can expect performance measures to evolve and improve as the focus of interventions changes (e.g., the decline in lead poisoning as the inventory of homes with lead-based paint drops) and as new public health threats are identified.

With this framework for a multilevel, user-oriented national information network in mind, the discussion now turns to steps that can be taken to promote the development of such a network. In discussing these steps, the panel presents recommendations both to help overcome barriers and to take advantage of opportunities. These recommendations address four broad concerns: policy actions to promote a collaborative approach to health-related performance measurement activities, operational principles to guide performance measurement, required investment in data systems and in training and technical assistance, and a national research agenda to improve the knowledge base for performance measurement.

MAJOR POLICY ACTIONS NEEDED

National Collaboration

The promotion of effective collaboration among multiple partners in the development of plans for performance measurement, the assessment of data needs, and the design (or redesign) of data systems to facilitate performance measurement must be a top priority. For example, focusing only on data to satisfy states' federal reporting requirements risks neglecting states' need for related but more detailed performance data that can be used in managing their program activities. Similarly, focusing unduly on individual categorical programs risks encouraging duplication of effort across programs, incompatibility across programs that rely on a program-specific data collection process, or lack of attention to essential functions (e.g., elements of the public health infrastructure) that are not specifically linked to individual programs.

Although the proposal to transform significant portions of DHHS grant funding to states into PPGs specifically addressed a federal-state relationship, performance measurement requires a much broader collaborative partnership across federal, state, and local governments if it is to succeed. The panel sees opportunities to strengthen health-related performance measurement efforts through collaborative efforts that would enable all parties to learn from each other and would promote consensus regarding health outcome, process, and capacity measures that are appropriate at the federal, state, and local levels. These efforts must

entail a sharing of responsibilities, as well as opportunities for participation in decision making.

Collaboration is needed not only across levels of government but also within each level. For example, more coordination among federal agencies in determining data needs would reduce duplication of effort at both the federal and state levels. Similarly, the more that states work together—in multistate regions, for example—and the more that agencies within the same state collaborate, the more likely they will be to learn from each other, share data, and develop comparable measures and definitions for data related to health outcomes and program activities.

Although the panel has focused primarily on the public-sector participants in performance monitoring, many private-sector players should also be considered partners in the process to one degree or another. For example, organizations that represent state and local public health officials and related professionals are playing an important role in planning and implementing performance monitoring. Health care institutions and organizations have a stake in decisions that affect data reporting requirements. Managed care plans should have a strong interest in performance goals that focus on reducing acute and chronic disease risks in populations they serve (e.g., increasing the percentage of older adults who receive influenza vaccinations; increasing the percentage of persons with diabetes who have had their blood glucose control checked within the past 12 months). Attention should be given to including these private-sector partners, as appropriate, in the collaborative performance measurement and monitoring process. One model might be the broad community-level public-private partnership for performance monitoring described by the Institute of Medicine (1997b).

To achieve the collaboration that the panel views as essential for effective performance measurement, the federal, state, and local partners in this effort should take the following recommended steps.

1. Federal, state, and local governments should commit to a common and national strategic goal of incorporating performance measurement into the practices of publicly funded health programs.

Achieving the goals of performance measurement requires a public and collaborative commitment by partners at the federal, state, and local levels and a strategic planning process for implementation. The performance measurement concept will also require ongoing advocacy and promotion. All levels of government should identify and use opportunities for dissemination of data that will encourage the use of the data for performance measurement. Performance measures and support for the development of integrated data systems that can facilitate the application of those measures should be incorporated in strategic planning documents with national scope, such as the Healthy People 2010 report

currently in preparation. Where appropriate, these public-sector efforts should interact with related activities in the private sector.

2. Federal, state, and local governments, with input from private partner organizations, should plan and implement all steps of the performance measurement process in full collaboration with one another.

A collaborative decision-making process will help ensure subsequent support for processes and investment in outcomes at all levels of government. The PPG proposal explicitly recognized states as partners in the performance measurement process, but a stronger basis for a broader, ongoing, and longer-term dialogue is needed. Collaborative efforts must extend across the federal, state, and local levels and will require the participation of a multiplicity of stakeholders from each level to ensure that all important issues are addressed.

Discussions at the national level among organizations representing the range of officials who will need to participate in performance monitoring efforts can help promote a shared national vision of and commitment to performance monitoring. The participants in such discussions should represent a mix of policy, program, and technical perspectives. Some of the relevant organizations are listed in Box 5-1.

An example of this type of collaboration is the new National Public Health Performance Standards Program (see Halverson et al., 1998). This partnership involves the CDC, the National Association of County and City Health Officials (NACCHO), the Association of State and Territorial Health Officials (ASTHO), the National Association of Local Boards of Health, the Public Health Foundation, and the American Public Health Association. The Public Health Performance Standards Program is an effort to define and measure core public health performance at the local and state levels. Measures of local public health performance will be incorporated into a new version of the Assessment Protocol for Excellence in Public Health (APEX/PH) that NACCHO is developing to assist strategic planning efforts by local public health systems throughout the country.

To facilitate such discussions and appropriate participation in decision making, a process is needed to bring together appropriate stakeholders in performance measurement. To accomplish this, the panel makes the following recommendation.

3. DHHS should work in partnership with members of the relevant groups representing policy, program, and technical officials of states and local entities to establish a process for developing policies and procedures that can facilitate the implementation of performance measurement efforts in health-related areas.

Box 5-1
**Examples of National Organizations That Might Represent
State and Local Perspectives in Discussions on
Performance Measurement**

Listed below are examples of national organizations representing state and local policy, program, and technical perspectives that should participate in discussions regarding health-related performance measurement activities. The panel emphasizes that this is an illustrative list and should not be considered comprehensive or definitive.

American Public Human Services Association
Association of Maternal and Child Health Programs
Association of Public Health Laboratories
Association of State and Territorial Health Officials
Association of State and Territorial Chronic Disease Program Directors
Council of State and Territorial Epidemiologists
Environmental Council of the States
National Alliance of State and Territorial AIDS Directors
National Association for Public Health Statistics and Information Systems
National Association of Counties
National Association of County and City Health Officials
National Association of Health Data Organizations
National Association of Local Boards of Health
National Association of State Alcohol and Drug Abuse Directors
National Association of State Emergency Medical Services Directors
National Association of State Medicaid Directors
National Association of State Mental Health Program Directors
National Conference of State Legislators
National Governors' Association
State and Territorial Injury Prevention Directors Association

Steps must be taken to initiate the consensus-building discussions that are needed to further the development of performance measurement efforts. A process for continuing these discussions must also be established. DHHS will be a key participant and may be an essential catalyst for this process, but must act as a partner with state and local stakeholders. Because many points of view must be considered fully and fairly, one approach might be to identify an interested party without a direct stake in the outcomes (e.g., a foundation, a university, a unit of the National Academy of Sciences) that can convene local, state, and national stakeholders in a neutral setting. Well-defined mechanisms should be established

for designating representatives of interested stakeholder groups to ensure that the views presented are authoritative and that channels exist for communicating with the stakeholder communities. A forum such as this might become a nationally recognized body through which many of this panel's recommendations might be advanced.

Consultation among federal, state, and local program representatives can help ensure that performance measures are consistent with priorities for health programs across levels of government. For example, federal partners should gain a better understanding of attributes of data systems that are critical to states and others. Without this input, such systems are not likely to address state-specific issues and concerns, and opportunities to link or compare national surveys with related state surveys may be missed. The National Health Interview Survey, for example, produces national estimates but has not routinely had input from state representatives. Thus while this survey meets the needs of many federal programs, it is currently of little direct use to state health programs.

The federal partners can also make important contributions to states and local entities in the planning and maintenance of performance monitoring systems. From their national vantage point, federal partners are often able to showcase information about a wider range of state activities than individual states could ascertain on their own. By serving as a clearinghouse for such information, the federal partners can disseminate problem-solving strategies developed in one state to other participants who might benefit from the information. Federal partners are also in a position to provide expertise and technical support to states (as was the case with the federal-state cooperation that contributed to the development of the national vital statistics system), and by serving as coordinating agents, they can promote comparability and quality through standardization of measures and data.

A process that allows states and communities to learn from each other can be expected to aid them in making more effective use of the resources available to support performance measurement. As an extension of a broad national discussion, a similar process should be established to help key executive agencies at the state and local levels (e.g., health, human services, education, natural resources) share the vision on which performance measurement is based, develop the administrative procedures needed to ensure appropriate and timely reporting, and provide the leadership required to encourage appropriate information sharing and data integration.

4. Federal, state, and local governments should accept explicit responsibilities, determined in collaboration with other stakeholders, in return for their share in the governance of and benefits from broader efforts to improve performance monitoring.

A well-designed and effectively operating performance monitoring system offers benefits for all of its participants. To achieve those benefits, however, participants must also accept an appropriate share of responsibility for the design and maintenance of the system and for an investment of resources so that sufficient capacity is available for the system to be used effectively. Among these responsibilities is the provision of adequate staff, facilities, and technical competence (e.g., in epidemiology, computer programming) to meet the objectives and standards of the data system. Moreover, participants at all levels must expect to bear a fair share of the financial cost in proportion to the degree to which the system meets their specific needs. All participants have a responsibility to work for compromise solutions in such matters as uniformity in definitions and procedures, choice of data items and data collection methods, and timeliness and format of the data forwarded to other collaborators.

Integration of Data Systems

Some states are beginning to pursue a more integrated approach to health data (U.S. Department of Health and Human Services, 1998), especially data related to health care services. However, the categorical nature of much of the federal funding for state and local health-related programs has often encouraged both a fragmented approach to health problems and the development of program-specific data systems and reporting requirements. Even though programmatic funding streams are likely to remain a prominent feature of federal funding, additional opportunities are needed at the state and local levels to improve data systems by coordinating and integrating a broader array of health data. Specifically, the panel makes the following recommendations.

5. DHHS should lead efforts to integrate data systems across categorical health program lines.

Coordination and integration of data systems across program areas promises increased quality, efficiency, timeliness, and usefulness for performance measurement and other purposes. As noted above, however, the categorical nature of federal funding streams and related reporting requirements often limits the opportunity for and value of such approaches. DHHS should develop specific incentives to encourage programs with separate categorical funding streams to develop integrated data systems. For example, federal requirements for the collection and transmission of data on HIV, tuberculosis, sexually transmitted diseases, and vaccine-preventable diseases should be mutually compatible and sufficiently flexible to ensure that states can easily collect and transmit the data using a single notifiable disease reporting system.

There is evidence that some federal agencies are recognizing the limitations

of the categorical approach that has to date shaped information system funding and development. CDC and HRSA have endorsed the use of their categorical grant funds in the development of integrated health information systems to derive the benefits of more comprehensive measures of health status, greater efficiency in managing programs, and more timely information for decision making (Broome and Fox, 1998). ASTHO and NACCHO have also endorsed this policy. The panel encourages all DHHS agencies that fund program activities at the state and local levels to promote this broader perspective in planning information system changes and to facilitate the use of funds from federal programs to implement those changes.

CDC and HRSA note, however, that their new policy does not supersede current legal restrictions limiting the use of grant funds for planning and evaluation or administrative expenses. The Mental Health Block Grant, which is administered by the Substance Abuse and Mental Health Services Administration (SAMHSA), limits states to using no more than 5 percent of these funds for such administrative functions as information system development and operation. To facilitate the implementation of recommendation 5 above, the panel also makes the following recommendation.

6. DHHS, in collaboration with state and local partners, should review restrictions on the use of grant funds to determine whether they represent a significant barrier to progress in the development of integrated health information systems. If so, DHHS should pursue changes in the terms of those grant programs that would permit greater flexibility in the use of the funds.

The review of any grant program should involve representative stakeholders from the national, state, and local levels.

Technology Policy

The rapidly evolving information technologies that will enhance the feasibility of performance measurement and the health information network envisioned by the panel (see Chapter 4) are also creating a need for standards that will facilitate the transmission and aggregation of data from multiple sources without requiring the use of specific equipment or operating systems and software. The development of capabilities for linking and merging electronic data from disparate sources (e.g., administrative records, patient records, vital statistics records, disease registries, disease surveillance systems, periodic surveys, needs assessment studies, social indicator systems) will remain important. The Internet and the World Wide Web offer increasing opportunities to transmit and use data easily and rapidly (see Chapter 4). Furthermore, as the cost of desktop computers drops and their capabilities increase, users can perform more complex analyses,

although many may have limited training to ensure that those analyses are appropriate (see the discussion of training needs below).

While the panel sees great opportunities in these technologies, it also sees a need for explicit efforts to oversee and guide their application to performance measurement to ensure their appropriate and effective use. Specifically, the panel makes the following recommendation.

7. DHHS should provide leadership in the development and use of data transmission standards and of new information technologies to collect, analyze, and disseminate health-related data.

The information revolution is proceeding at such a rapid pace that many state and local health agencies cannot remain current. In addition, standardization is a key feature of successful integration and interoperability of data systems. DHHS should serve as a catalyst for consensus building on standards for data collection and transmission and as a technical assistance resource for jurisdictions seeking help and guidance on the appropriate application of new information technologies for a broad range of health-related data. The standards development process mandated by the Health Insurance Portability and Accountability Act (HIPAA) is achieving substantial progress for administrative transactions of health plans; however, these standards are not likely to be sufficient for performance measurement because they are not intended to address data and data systems for the population-based services provided by most state and local health departments. As standards are developed, attention should also be given to opportunities to facilitate the linkage of data sets. Moreover, all activities related to the development and use of information technologies must address the protection of individual privacy and assurance of the confidentiality and security of health-related data (see Chapter 4).

OPERATIONAL PRINCIPLES

In addition to considering policy actions required for the implementation of an effective performance measurement information system, the panel has developed recommendations regarding three operational principles that it believes are important for the success of such a system: involve a broad range of agencies, build on existing information systems, and provide for ongoing review and standardization of performance measures and data systems.

Involve a Broad Range of Public Agencies

8. As states and communities work to implement performance monitoring systems for health-related programs, they should ensure that

all relevant public agencies, including those outside traditional health areas, have the opportunity to participate.

The development of a performance monitoring system requires a commitment to understanding the range of substantive factors that influence desired outcomes for health and well-being, as well as the administrative, analytic, and technical resources needed to collect and use performance measures. The process must ensure participation not only by state and local health agencies traditionally responsible for the health of the public, but also those with programmatic responsibilities in relevant nonhealth areas (e.g., criminal justice, housing, transportation) and those that collect and manage data on basic socioeconomic characteristics of the population and the state or community (e.g., population estimates, economic development data). A community health improvement process embracing this broad involvement in performance measurement has been described by the Institute of Medicine (1997b), and the implications of such an approach for population information systems are explored by Roos and colleagues (1995).

Build on Existing Information Systems

A health information network should facilitate access to all available information on populations that could be used to track health risks and promote the health of communities. Substantial amounts of health-related data are currently being collected through existing data systems at the federal, state, and local levels. Many states and localities have the capability for and are already invested in the collection of performance data because these data comprise a subset of the information needed to operate efficient and effective programs. Although few existing data systems have been designed specifically for performance measurement, they nonetheless provide an essential base from which to build an information network that can meet a broad range of health information needs, including performance data. Use of existing data systems to the extent feasible would provide an important efficiency in the establishment of a performance measurement system. Thus, the panel makes the following recommendation.

9. When possible, partners should obtain performance measurement information from existing or enhanced federal, state, and local information systems.

Instead of creating an entirely new data system for the purpose of performance measurement, data needs should be met to the extent possible through the use of existing systems. Systems such as vital records, the notifiable disease systems, adult and youth behavioral risk factor surveillance, cancer registries, and immunization registries can provide data for national, state, and local estimates regarding health risks, disease incidence and prevalence, treatment, and

health outcomes and should serve as the foundation for performance measurement.

When existing data systems cannot meet information needs for performance measurement, investments must be made to enhance those systems or, in cases where enhancement is not feasible or sufficient, to develop new data systems. These investments can potentially yield multiple benefits across programs and across national, state, and local lines, but modifications should reflect a careful balancing of costs and benefits. For example, enhanced technical assistance and improved standardization of BRFSS methodology may involve costs and loss of some flexibility by states, but it may also produce more directly comparable state-specific estimates and data that could be aggregated to yield meaningful national estimates. The panel supports efforts to make the data systems in a health information network as efficient and effective as possible.

As noted in conjunction with recommendation 8, an information network should also provide access to data systems that are beyond the purview of health agencies but contain data valuable for tracking health risks and outcomes. For example, law enforcement and transportation department data on traffic accident locations and times have been used to identify high-risk intersections and highways. By linking these data with emergency medical services records and hospital emergency room data, a health department may be able to distinguish physical risks in the community (e.g., poorly marked curves and intersections) from behavioral risks (e.g., times of day when alcohol-related injuries are more likely to occur) (Perrin et al., 1996). Similarly, access to law enforcement data on drug-related crimes is crucial for monitoring the effectiveness of some substance abuse treatment services.

Implementation Standards

Performance measures themselves require further development and standardization. For some program areas, a basic set of outcome measures is not yet recognized. For others, the evidence base for capacity, process, and risk status measures is still developing. In most program areas, understanding of the relationship between measures that make sense at the state or local level and national measures is still limited. Furthermore, health needs and program priorities can be expected to change over time. Thus, the selection and review of performance measures and the development of related standards must be a continuing effort that takes into account the evolving knowledge base on which such measures rest, changing needs for and opportunities to obtain data, and the changing program environment in which performance measures will be used.

Once suitable measures have been identified, the adoption of standard definitions and procedures for data collection will enhance the comparability of performance data over time and across states and localities. One of the challenges for the health information network envisioned by the panel will be balancing flexibil-

ity against standardization. Flexibility is needed to accommodate a broad range of measures as they continue to evolve. Standardization is required to produce data and measures that permit comparisons among communities, states, and regions, and that allow data to be tracked over time or pooled for regional and national estimates. To promote the continued development and improved comparability of performance measures, the panel recommends the following steps.

10. DHHS, in partnership with state and local stakeholders, should lead the implementation of a process for ongoing development and review of performance measures to be used in conjunction with state and local health programs.

The recommended process should be a collaborative effort that includes participation by federal, state, and local health agencies, plus representatives of consumers and relevant private-sector interests (e.g., health plans, clinicians, insurers, businesses, foundations, patient advocates, grass roots organizations, and the general public). Involvement of a broad range of stakeholders is recommended to ensure that performance measures are consistent with state and local public health priorities and that policy, programmatic, and technical perspectives are all represented. The ongoing collaboration between CDC and the Council of State and Territorial Epidemiologists for periodic review and revision of the case definitions of specific infectious diseases might serve as a model for these efforts (Centers for Disease Control and Prevention, 1997).

This process must also take into account the differing stages of development among various health fields. For example, the mental health field should first emphasize the development of a standardized framework for the evaluation of mental health services that reflects an understanding of the relationship of outcomes to program capacity and processes. In other fields, such as immunization, the framework for assessing program activities is more fully developed. For those fields in which greater progress has been made in defining outcome measures, the process should focus on reviewing and refining existing performance measures to ensure that they reflect current knowledge regarding causal relationships between outcomes and processes and capacity.

In some fields (e.g., substance abuse and mental health), the development and review of performance measures must be preceded by efforts to build consensus on a framework for assessing health outcomes. The panel's earlier efforts to identify performance measures in these fields were hindered by the current diversity of evaluation systems among states, provider organizations, and accrediting bodies and by the resulting inconsistencies in measurement and data collection. A broadly based effort will be needed to bring together relevant federal, state, and local agencies, professional groups, and consumer interests to seek consensus on the dimensions of health outcomes, processes, and capacities that should serve as the basis for performance assessment in these fields. Chapter 3

notes some of the work that has been initiated in the fields of mental health and substance abuse.

The recommended review process should also examine both the positive and negative effects that the use of these performance measures in a performance monitoring system might have (or might be expected to have) on program activities and outcomes. Positive effects would be anticipated, and should be verified, for those measures that encourage programs to implement evidence-based best practices. Other measures, however, might conceivably result in a negative impact if, for example, they encouraged restrictions in program services for hard-to-reach or high-risk groups to achieve the appearance of better performance.

11. DHHS, in partnership with state and local stakeholders, should lead a process for assembling and evaluating sets of performance measures from which users can identify and agree upon those appropriate for specific applications.

Because of data limitations, differing health problems, and differing priorities, users of performance measures need to select a suitable subset of health outcome measures from a larger array of measures that have been found to meet the basic tests of validity, reliability, responsiveness, and data adequacy. States and localities may reasonably pursue many different strategies to target a single health outcome, so users should have an even larger number of process and capacity measures associated with these outcomes from which to choose. Efforts by private nonprofit groups to identify measures suitable for outcomes research in health care (e.g., the Medical Outcomes Trust) might serve as models for a process for assembling sets of performance measures for publicly funded health programs.

The panel's first report (National Research Council, 1997) provided examples of measures of health outcomes, program processes, and capacities that could be used to monitor performance in specific program areas. The panel concluded that it could not propose a definitive list of measures because few states are likely to have the data necessary to support every measure, and individual agencies are likely to have priorities in addition to (or different from) those reflected in the measures selected by the panel.

Instead, one of the panel's principal goals in its first report was to provide an analytic framework that could be used to assess the appropriateness of specific outcome, process, and capacity measures proposed for performance monitoring purposes, and could be applied over time to modify or replace measures as new or better ones were identified. Using this framework, it would be possible to develop a menu from which suitable measures could be chosen. An effort should be made, however, to associate particular program goals and strategies with specific outcome, risk status, process, and capacity measures so that identical activities related to those goals and strategies can be monitored using the same measures.

12. DHHS should work in partnership with state and local stakeholders to promote the development and adoption of standard definitions for performance measures and standards for associated data collection and data quality in performance measurement systems.

Comparability of Data and Data Collection Methods While specific applications of performance measurement may vary, common definitions are critical for any communication and comparison of results. Similarly, while various data collection methods may be used to obtain performance data, standards for each method should be promoted in all cases to enhance data quality and promote comparability.

For performance measures, agreement is needed on the definitions of the measures and on the data elements to be used to construct the measures. For a measure of adolescent tobacco use, for example, the definitions must cover age (e.g., 14-17 years, 13-18 years, 12-17 years), the scope of tobacco use (e.g., smoking cigarettes, all smoked tobacco, smoked and chewed tobacco), and the time reference (e.g., the past week, the past month, the past year). The choices made should reflect considerations of scientific and policy relevance and of practical factors such as data availability.

Consideration should also be given to specifying the most appropriate method of data collection. Many data collection methods can be used to produce performance data. Potential methods include sample surveys, which can employ techniques such as mailed questionnaires, telephone interviews, and in-person interviews; clinical or administrative records from service encounters; reports to surveillance systems; and environmental monitoring. The data collection process cannot be specified in complete detail, but essential guidelines can be established so that differences in data collection methods will not undermine the comparability gained through the use of standard definitions for performance measures.

If standards for performance measures and data collection are to be developed and accepted, a broad range of stakeholders must have a means of achieving consensus and harmonizing the implementation of their performance measurement activities. For example, DHHS has worked with the states to reach agreement on common definitions and practices for vital registration that permit the valid compilation of state data to produce national figures. Similar collaborative efforts in other areas should be able to produce sufficient comparability across state data systems so that greater reliance could be placed on aggregating state data to produce national measures, rather than requiring separate data collection systems at the federal level.

Consensus-building mechanisms also exist in the private sector, as illustrated by the participation of many organizations and individuals in voluntary standards-setting activities, such as efforts under the auspices of the American Society for Testing and Materials (ASTM) to develop data and systems standards for certain types of clinical information. Another example is the Performance Measurement

Coordinating Council, formed in 1998 among the American Medical Accreditation Program, the Joint Commission on Accreditation of Healthcare Organizations, and the National Committee for Quality Assurance to coordinate performance measurement activities across the health care system (Joint Commission on Accreditation of Healthcare Organizations, 1998). The American National Standards Institute (ANSI) is another useful mechanism. It accredits a variety of consensus-based voluntary standards-setting activities and helps promulgate and maintain the standards that are adopted through those activities.

If more jurisdictions used common measures for performance monitoring, greater numbers of valid comparisons could be made among those jurisdictions. Use of the same measures and data definitions would also encourage cost efficiencies by reducing the need to redesign data collection instruments, electronic processing protocols, and similar infrastructure elements. The advantages of this harmonization of measures and data collection practices for comparability and efficiency should not, however, be allowed to obscure the need for continued critical assessment of their appropriateness in general or for specific applications.

Data Quality Data systems that support performance measurement must ensure adequate quality and appropriate handling of the data that are collected. Data quality can be compromised in many ways, such as inaccurate or incomplete reporting, poorly designed survey samples (including nonresponse), errors introduced in data processing procedures, and inaccurate calculation of measures. In addition to consideration of such problems within individual data systems, a broader information network such as that envisioned by the panel would require consideration of the differences in data quality across information systems and their impact on comparisons of performance measures.

No data system or data set is ever perfect, and costs tend to rise rapidly as residual errors are further reduced. An appropriate balance must be found between the desired degree of data quality and the cost of achieving it. The requirements for data quality should be judged largely on the basis of the intended and anticipated uses of the data, but with some consideration of future uses not yet foreseen. At the same time, data systems should not be overdesigned in an attempt to anticipate all possible future uses. For example, efforts by cancer registries to collect accurate data on treatment have often been scaled back because the necessary quality and uniformity could not be attained on a population-wide basis.

The effects on data quality of bias (e.g., survey response rates that differ across population groups) and random variation should be considered separately because these two factors have largely different sources and different implications. For example, bias that is constant over time may have little impact if policy considerations focus on time trends. Similarly, in situations where bias dominates random variation, there may be little profit in further reducing the random component of uncertainty. Increasing the size of a survey's sample, for example,

would tend to reduce random error but would not overcome the bias introduced by a factor such as widespread underreporting of tobacco use.

Standards for data quality and the methods adopted to meet those standards should be based on informed assessments of how the data are to be used and the degree of accuracy and precision needed to serve those uses. The panel also advises consideration of policies and practices that will help achieve and maintain the desired level of data quality. Many observers agree that making data useful to those who produce them is one of the stronger incentives for ensuring that the data are of high quality.

REQUIRED INVESTMENT FOR AN INFORMATION NETWORK

Performance measurement activities can be expected to impose new demands (e.g., additional data collection, data system development, and analysis of performance data) on those whose performance is being assessed. The panel sees a need to ensure that adequate resources, including funding, trained staff, and technology, are available to meet those demands, as well as to maintain the effective elements of current data systems. Moreover, because resources are often so limited, the panel is concerned that information system development will be forced to compete for funds with program services. To respond to these concerns, the panel recommends investments both in data systems and in training and technical assistance for health agency personnel, who will be expected to assume responsibility for planning and implementing performance measurement.

Data Systems

Existing health data systems provide a strong base for performance measurement, but because they have generally not been developed for this purpose, information gaps will exist. To address these gaps, the panel makes the following recommendation.

13. DHHS and state and local users of performance measurement data should each commit resources to reduce gaps in the supporting information systems.

Adequate resources are needed to maintain key information systems that provide essential data for performance measurement and to enhance or develop new systems for data that cannot currently be produced. DHHS acknowledges that obtaining better data for performance measurement, especially state-level data, will require investments in data systems at the federal and state levels (U.S. Department of Health and Human Services, 1997).

The panel urges specific attention to the need for resources to ensure and improve the availability of data for state- and local-level performance measures.

For example, many of the measures identified in the panel's first report depend on data produced by the BRFSS, for which a core survey is fielded in every state. Overall, grants from CDC's BRFSS program support about half of this collaborative state-CDC data system. Recent instability in CDC funding is of concern to the panel. Funding of \$3.5 million for the 1996-1997 grant cycle was reduced by nearly one-quarter, to \$2.7 million, for the 1997-1998 cycle (D. Nelson, Centers for Disease Control and Prevention, personal communication, 1998). An increase to \$3.9 million for the 1998-1999 grants is encouraging, but the adequacy of the level and stability of federal support for this program should be assessed. A positive sign is the Maternal and Child Health Bureau's State Systems Development Initiative, which specifically allows states to apply for grants of up to \$100,000 that can be used for such purposes as the development of integrated information systems or the collection of data for performance measures for the Maternal and Child Health Services Block Grant (Maternal and Child Health Bureau, 1998).

Other important sources of state and local data that require adequate support include public health surveillance systems, such as vital records, notifiable disease systems, and registries for cancers, immunizations, and birth defects. Administrative and clinical data systems such as those that track client services in mental health and substance abuse programs should be assessed to determine their potential to contribute data at this level of geographic detail. The panel specifically noted in its first report that the mental health and substance abuse fields will have to develop ongoing information systems to provide most of the data necessary for performance measurement. SAMHSA's proposed Data Infrastructure Development funding should be helpful in this regard.

The federal and state agencies that are establishing performance measurement requirements should support the development and operation of the information systems needed to produce the required performance data. Because the health information network envisioned by the panel would have the potential to produce data of value to a variety of audiences in the public and private sectors, those data users might provide additional resources.

At the same time, the panel recognizes that new funding for strengthening and maintaining a health information network may be difficult to obtain and does not want to see funding for services compromised to support information systems. Therefore, innovative ways of using existing resources are needed. Some resources currently invested in data collection and analysis activities at the federal, state, and local levels could perhaps be redirected to produce more useful data. To this end, the panel recommends the following step.

14. DHHS should sponsor a review of the current array of federal, state, and local data collection and analysis activities to begin an assessment of how existing resources might be used most effectively to meet performance measurement and other needs for health data.

This review must include participation by appropriate state and local representatives.

DHHS efforts to review and better coordinate some of its major national health surveys (U.S. Department of Health and Human Services, 1995) are a good start, but a broader approach will be needed. The recommended review of data collection and analysis activities must take into account data needs, interests, and resources at the local, state, and federal levels, and may benefit from private-sector participation. Duplication in current data collection and data systems should be examined carefully to determine whether information needs can be met more efficiently, potentially freeing resources to improve or expand weaker data systems. All participants in this effort will have some stake in existing information systems and may be reluctant to accept changes in those systems. However, successful development of a broader national health information network that can support performance measurement and other operational, managerial, and analytic activities will require coordination and compromise, as well as a broad, long-term commitment by the participants.

The Environmental Protection Agency's (1998) One Stop Reporting Program, through which the agency is working with states to improve all aspects of environmental reporting, might serve as a model for a comprehensive assessment of health information systems. The panel notes that this program also allows states that are investing in environmental information system improvements to apply for demonstration grants of \$500,000 to support such efforts.

Training and Technical Assistance

To have an effective information system for performance measurement, it will be essential to invest not only in the processes and equipment needed for data collection and analysis but also in enhancement of the knowledge and expertise of the information system staff and others who will produce and use performance data. The panel makes the following recommendations with regard to training and technical assistance.

15. To ensure the success of performance measurement, all stakeholders, with substantial leadership from DHHS, should contribute ongoing technical assistance, training, and resources to enhance state and local data systems and analytic capacity.

To make effective performance measurement a reality, DHHS must actively support efforts to increase competency in analytic, programmatic, and leadership skills at the federal, state, and local levels. Resources must be allocated to fund and train staff at various levels, not just in the collection and analysis of perfor-

mance data but also in use of the results to bring about improvements in health programs.

Throughout the nation, states and communities are assuming more responsibility for their own health through priority setting, development and implementation of programs, and performance measurement. As a consequence, there is a growing need for state and community capacity for data collection and analysis. Staff vary tremendously, however, in their knowledge of the relevant disciplines and methodologies (e.g., epidemiology, statistics, social science research) and in their experience with the use of data to plan, evaluate, and revise community programs. Moreover, state and local health agencies are often understaffed, making it difficult to take on the additional tasks in data collection and analysis that performance measurement may entail.

A variety of approaches might help in capacity building. A useful first step might be facilitating access to the funds and expertise required to assess specific capacity needs. One community or state might have a pressing need for data processing resources (e.g., expertise and equipment), while another might need expertise in survey design or data analysis. Because of these differing needs, varying means of meeting them will be necessary.

CDC's Public Health Training Network might be one mechanism, or model, through which training materials could be developed and delivered to state and local audiences. Distance-based learning programs, which reduce the need to travel by establishing video and audio connections to multiple sites, should be considered as a way of reaching this broad audience. Additional funding specifically for travel to major conferences and training programs could also be an effective investment. Direct interaction with colleagues provides learning opportunities that are not available with other approaches, such as use of self-guided instruction or distance-based learning. Access to scholarships and dissertation grants could assist staff in obtaining more advanced academic training.

Other opportunities for state and local training and capacity building should also be explored. For example, university medical centers and schools of public health could develop collaborative programs that would provide state and local health departments with access to training and academic expertise in data collection and analysis while providing faculty and students with opportunities for field experience and research. Foundations, health plans, businesses, and others with interest or experience in performance measurement might become partners with state and local health agencies that are working to improve their skills.

16. DHHS should develop and maintain information technology expertise to assist states and communities as they use new technologies to improve the quality of and capacity for data collection, analysis, and dissemination.

Effective use of information technology requires considerable expertise in matters such as hardware, software, systems design and integration, and applications development. To the extent that DHHS can serve as a central and authoritative source for information and guidance on these technology matters, state and local health agencies will be able to make more rapid use of a broader range of expertise than they would be likely to assemble on their own. In the current highly competitive market for information technology personnel, many state and local agencies are at a disadvantage because their salaries tend to be relatively low.

One example of work DHHS is already doing along these lines is the preparation of an investment analysis guide for states that want to combine categorical grant funds for the development of integrated health information systems (Centers for Disease Control and Prevention and Health Resources and Services Administration, 1998). Another specific area for support is exploration of the technical issues involved in merging data from diverse systems. Among the activities DHHS might support are pilot projects to test the use of specific equipment or processes. DHHS might also promote the development of a national clearinghouse or similar repository for information system applications and modular programming components (i.e., software objects). Access to such resources can make the systems development process simpler, faster, and less costly.

A NATIONAL RESEARCH AGENDA

The panel emphasizes that a multidisciplinary research program must be an integral part of any ongoing performance measurement activity for health-related programs. Because experience with performance measurement is still limited, studies are needed to improve understanding of what measures and methods of data collection are appropriate. Also crucial is further research to establish evidence regarding causal links between program interventions and desired outcomes. This evidence, which is currently limited in many fields, is essential for selecting demonstrably meaningful capacity, process, and risk status measures. For example, recent studies have shown that the identification and treatment of co-occurring mental health problems are an important predictor of positive outcomes in substance abuse treatment (Harrison and Asche, in press). Specifically, the panel makes the following recommendation.

17. Federal agencies, foundations, and other private-sector groups should develop and fund a research agenda to support performance measurement activities, including the testing of intervention effectiveness, the investigation of the links between program capacity and processes and program outcomes, the development of measures, the

refinement of data collection and information system technologies, and the use of performance measurement systems and performance-based decision making.

Most users of performance measures will need to rely on the research of others to identify effective health interventions and measures of outcomes for these interventions, as well as to determine the effective use of performance measurement as a management and oversight tool. Therefore, the agendas of federal agencies and private organizations that fund health-related programs should include support for research intended to inform performance measurement in these areas. The range of studies must be broad enough to ensure that they are informative for a variety of settings, including the local, state, and national levels. Initiatives in the Agency for Health Care Policy and Research (AHCPR) under the designations of Strengthen Quality Measurement and Improvement and Support Improvements in Health Outcomes are models for the types of federal research agendas that could serve well for performance measurement research in other health areas.

However, carrying out the research is not, in itself, sufficient for informing and improving the performance monitoring process. Significant findings must be communicated to the participating partners before they can be applied. Therefore, the panel offers this further recommendation.

18. DHHS, foundations and other private organizations, and other partners involved in performance measurement activities should contribute in an appropriate manner to a process of information gathering and dissemination to support the use of evidence-based performance measures.

Resources are needed to support comprehensive reviews and rigorous analyses of the relevant scientific evidence and to produce evidence-based reports describing the scientific foundation applicable to the development of performance measures for use in publicly funded health programs. AHCPR currently sponsors a program to develop, use, and evaluate evidence-based tools and information related to clinical health care. A similar program that would apply to the areas of public health, mental health, and substance abuse would be immensely valuable for developing and refining performance measures and improving the performance monitoring system. Although states and communities generally have limited funds for conducting research or disseminating research results, they can make an important nonmonetary contribution to these efforts by facilitating studies undertaken by others and sharing information about their experiences in the use of performance measurement.

FINAL OBSERVATIONS

The development and implementation of a broad national health information network that can support performance measurement will require a concerted effort. The leadership of DHHS and its agencies is critical, but these federal agencies must be part of a collaborative undertaking that includes the participation of stakeholders at the state and local levels, in the public and private sectors, and across a variety of program areas. The participants must be partners in seeking benefits from and sharing responsibility for the effective management and operation of a broader health information network.

The information network envisioned by the panel should enhance the capacity of federal, state, and local health programs to meet performance measurement obligations and to use performance data and other information to achieve desired health outcomes. Several significant challenges lie ahead. Appropriate performance measures must be developed, and the data needed to use those measures must be available. Greater consensus must be achieved regarding standards for measures and data that will promote comparability in performance measurement. Necessary technical and analytic skills must be developed and applied to the creation and use of performance data. Sufficient financial and nonfinancial resources must be obtained to support both near-term efforts to introduce performance measurement activities and the longer-term data collection and analysis and research needed to sustain those efforts. This may appear to be a difficult set of challenges, but the panel believes that the need to address them is great and that the current commitment to performance measurement creates an opportunity to make significant progress toward meeting this need in the near future.

REFERENCES

- Broome, C.V., and C.E. Fox
1998 CDC/HRSA Grant Funding Flexibility for Integrated Health Information Systems. Grant funding transmittal letter. April 1, 1998. U.S. Department of Health and Human Services. <http://www.hrsa.dhhs.gov/policy.htm> (also at <http://www.cdc.gov/funds/policy.htm>) (April 21, 1998).
- Centers for Disease Control and Prevention
1997 Case definitions for infectious conditions under public health surveillance. *MMWR* 46(RR-10).
- Centers for Disease Control and Prevention and Health Resources and Services Administration
1998 Integrated Health Information Systems Investment Analysis Guide. <http://www.hrsa.dhhs.gov/investment.htm#iv> (also at <http://www.cdc.gov/funds/invest7.htm>) (April 21, 1998).
- Environmental Protection Agency
1998 One Stop Program Strategy and Grant Award Criteria. <http://www.epa.gov/reinvent/onestop/strategy.htm> (April 28, 1998).
- Gold, M.
1995 *Miss or Match: How Well Do State Data Systems Meet State Health Policy Needs?* Washington, D.C.: Mathematica Policy Research.

Halverson, P., R.M. Nicola, and E.L. Baker

- 1998 Performance measurement and accreditation of public health organizations: A call to action. *Journal of Public Health Management and Practice* 4(4):5-7.

Harrison, P.A., and S. Asche

- in press Comparison of substance abuse treatment outcomes for inpatients and outpatients. *Journal of Substance Abuse Treatment*.

Institute of Medicine

- 1997a *The Computer-Based Patient Record: An Essential Technology for Health Care*, Revised ed. R.S. Dick, E.B. Steen, and D.E. Detmer, eds. Committee on Improving the Patient Record. Washington, D.C.: National Academy Press.
- 1997b *Improving Health in the Community: A Role for Performance Monitoring*. J.S. Durch, L.A. Bailey, and M.A. Stoto, eds. Committee on Using Performance Monitoring to Improve Community Health. Washington, D.C.: National Academy Press.

Joint Commission on Accreditation of Healthcare Organizations

- 1998 Nation's Three Leading Health Care Quality Oversight Bodies to Coordinate Measurement Activities. Press release. May 19, 1998. <http://www.jcaho.org/news/nb.htm> (June 5, 1998).

Lasker, R.D., B.L. Humphreys, and W.R. Braithwaite

- 1995 Making a Powerful Connection: The Health of the Public and the National Information Infrastructure. Report of the U.S. Public Health Service Public Health Data Policy Coordinating Committee. Washington, D.C. <http://www.nlm.nih.gov/pubs/staffpubs/lo/makingpd.html> (August 11, 1998).

Maternal and Child Health Bureau

- 1998 State Systems Development Initiative (SSDI) Grant Application Guidance for FY98. U.S. Department of Health and Human Services, Health Resources and Services Administration. <http://www.hhs.gov:80/hrsa/mchb/guidance.htm> (June 4, 1998).

Mendelson, D.N., and E.M. Salinsky

- 1997 Health information systems and the role of state government. *Health Affairs* 16(3):106-119.

National Center for Health Statistics

- 1998 National Vital Statistics System. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. <http://www.cdc.gov/nchswwww/about/major/nvss/nvss.htm> (July 6, 1998).

National Committee for Quality Assurance

- 1997 *HEDIS 3.0/1998*. Vol. 4, *A Roadmap for Information Systems: Evolving Systems to Support Performance Measurement*. Washington, D.C.: National Committee for Quality Assurance.

National Research Council

- 1997 *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health*. E.B. Perrin and J.J. Koshel, eds. Panel on Performance Measures and Data for Public Health Performance Partnership Grants, Committee on National Statistics. Washington, D.C.: National Academy Press.

Perrin, E.B., L.G. Hart, S.M. Skillman, B. Paul, M.A. Hanken, and J. Hummel

- 1996 *Health Information Systems and Their Role in Rural Health Services: Issues and Recommendations*. Report to the Agency for Health Care Policy and Research (Contract #282-93-0036-4). Seattle: University of Washington.

Public Health Foundation

- 1998 *Measuring Health Objectives and Indicators: 1997 State and Local Capacity Survey*. Washington, D.C.: Public Health Foundation.

Roos, N.P., C.D. Black, N. Frohlich, C. Decoster, M.M. Cohen, D.J. Tataryn, C.A. Mustard, F. Toll, K.C. Carriere, C.A. Burchill, et al.

- 1995 A population-based health information system. *Medical Care* 33(12 suppl.):DS13-DS20.

Starr, P.

- 1997 Smart technology, stunted policy: Developing health information networks. *Health Affairs* 16(3):91-105.

Thacker, S.B., and D.F. Stroup

- 1994 Future directions for comprehensive public health surveillance and health information systems in the United States. *American Journal of Epidemiology* 140:383-397.

U.S. Department of Health and Human Services

- 1995 HHS Plan for Consolidation of Surveys. April 11, 1995. Washington, D.C. <http://aspe.os.dhhs.gov/datacncl/srvyrpt1.htm> (January 12, 1998).
1997 1997 Strategic Plan. September 30, 1997. Washington, D.C.
1998 Registry of State-Level Efforts to Integrate Health Information. <http://aspe.os.dhhs.gov/statereg/index.htm> (February 9, 1998).

Appendices

APPENDIX A

Potential Health Outcome and Risk Status Measures

The health outcome and risk status measures in this appendix are presented to illustrate the types of measures that might be included in performance partnership grants (PPGs) between state agencies and the U.S. Department of Health and Human Services (DHHS). These measures were selected from among the many proposed to the panel by participants at four regional meetings sponsored by DHHS, as well as by professional health associations and private agencies and individuals. The panel chose the measures using the guidelines described in Chapter 1 of this report: a measure should be specific and results oriented; it should be meaningful and understandable; data should be adequate to support the measure; and the measure should be as valid, reliable, and responsive as possible.

These health outcome and risk status measures are not meant to represent a mandated list. Few states are likely to have all of the data necessary to support all of these measures. In addition, state agencies may well have major priorities beyond those represented by the categories of outcome measures listed here (e.g., injury prevention, oral health, hearing and vision, environmental health) and are responsible for administering major programs relevant to public health that are not covered by this report (e.g., Medicaid). In addition, the panel did not attempt to identify all of the measures that might be relevant for specific important subpopulations (i.e., groups defined by demographic or risk categories). Consequently, the health outcome and risk status measures shown below should be

NOTE: This appendix is adapted from Appendix C of this panel's first report, *Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health* (National Research Council, 1997).

considered an important subset, but not an exhaustive listing, of those that will be of interest to state agencies.

A major goal of this report is to provide an analytic framework for use by the states and DHHS in assessing the appropriateness of specific outcome, process, and capacity measures proposed for PPG agreements in the future. The panel hopes that the field of performance measure evaluation will evolve as new health outcome measures are defined and studied and become available. It is anticipated that many of the measures described in this report can, in time, be modified or replaced by others that meet the selection guidelines cited above.

POTENTIAL MEASURES: OVERVIEW

Chronic Disease

Tobacco

- Percentage of (a) persons aged 18-24 and (b) persons aged 25 and older currently smoking tobacco
- Percentage of persons aged 14-17 (grades 9-12) currently smoking tobacco
- Percentage of women who gave birth in the past year and reported smoking tobacco during pregnancy
- Percentage of employed adults whose workplace has an official policy that bans smoking

Nutrition

- Percentage of persons aged 18 and older who eat five or more servings of fruits and vegetables per day¹
- Percentage of persons aged 14-17 (grades 9-12) who eat five or more servings of fruits and vegetables per day²
- Percentage of persons aged 18 and older who are 20 percent or more above optimal body mass index³

Exercise

- Percentage of persons aged 18 and older who do not engage in physical activity or exercise

¹The numerical value in this measure is the level that is generally regarded as appropriate by the medical community; it does not represent a level that has been independently determined or endorsed by the panel.

²See fn. 1.

³See fn. 1.

- Percentage of persons aged 14-17 (grades 9-12) who do not engage in physical activity or exercise

Screenings and Tests

- Percentage of persons aged 18 and older who had their blood pressure checked within the past 2 years⁴
- Percentage of women aged 45 and older and men aged 35 and older who had their cholesterol checked within the past 5 years⁵
- Percentage of women aged 50 and older who received a mammogram within the past 2 years⁶
- Percentage of adults aged 50 and older who had a fecal occult blood test within the past 12 months or a flexible sigmoidoscopy within the past 5 years⁷
- Percentage of women aged 18 and older who received a Pap smear within the past 3 years⁸
- Percentage of persons with diabetes who had HbA1C checked within the past 12 months⁹
- Percentage of persons with diabetes who had a health professional examine their feet at least once within the past 12 months¹⁰
- Percentage of persons with diabetes who received a dilated eye exam within the past 12 months¹¹

STDs, HIV Infection, and Tuberculosis

- Incidence rates of selected STDs
- Incidence rates of HIV infection
- Prevalence rates of selected STDs
- Prevalence rates of HIV infection
- Consumer satisfaction with STD, HIV, and tuberculosis treatment programs
- Rates of sexual activity among adolescents aged 14-17
- Rates of sexual activity with multiple sex partners among people aged 18 and older
- Rates of condom use during last episode of sexual intercourse among sexually active adolescents aged 14-17

⁴See fn. 1.

⁵See fn. 1.

⁶Cancer incidence by diagnosed stage may be a better alternative in cancer registry areas; see fn. 1.

⁷See fns. 1 and 6.

⁸See fns. 1 and 6.

⁹See fn. 1.

¹⁰See fn. 1.

¹¹See fn. 1.

- Rates of condom use during last episode of sexual intercourse by persons aged 18 and older with multiple sex partners
- Rates of condom use during last episode of sexual intercourse among men having sex with men
- Rates of injection drug use among adolescents and adults
- Completion rates of treatment for STDs, HIV infection, and tuberculosis

Mental Health

- Percentage of persons aged 18 and older receiving mental health services who experience reduced psychological distress
- Percentage of persons aged 18 and older receiving mental health services who experience increased level of functioning
- Percentage of persons aged 18 and older receiving mental health services who report increased employment (including volunteer time)
- Percentage of persons aged 18 and older with serious and persistent mental illness receiving mental health services who live in integrated, independent living situations or with family members
- Percentage of children aged 17 and younger with serious emotional disorders receiving mental health services who live in noncustodial living situations
- Percentage of persons aged 18 and older with serious mental illness who are in prisons and jails
- Percentage of children aged 17 and younger with serious emotional disorders who are in juvenile justice facilities
- Percentage of homeless persons aged 18 and older who have a serious mental illness
- Percentage of adolescents aged 14-17 or family members of children and adolescents or both who are satisfied with: (a) access to services, (b) appropriateness of services, and (c) perceptions of gain in personal outcomes
- Percentage of persons (aged 18 and older) or their family members or both who are satisfied with: (a) access to mental health services, (b) appropriateness of services, and (c) perceptions of gain in personal outcomes

Immunization

- Reported incidence rate of representative vaccine-preventable diseases
- Age-appropriate vaccination rates for target age groups (children aged 2 years; children entering school at approximately 5 years of age; and adults aged 65 and older) for each major vaccine group

Substance Abuse

- Death rate of persons aged 15-65 attributed to (a) alcohol, (b) other drug use, and (c) combined agents

- Percentage of emergency room encounters for alcohol or other drug-related causes
- Prevalence rate of substance abuse clients who report experiencing diminished severity of problems after completing treatment as measured by the Addiction Severity Index (ASI) or a similar measure¹²
- Ratio of substance abuse clients involved with the criminal justice system before and after completing treatment
- Prevalence rate of adolescents aged 14-17 engaged in heavy drinking or other drug use¹³
- Prevalence rate of persons aged 18 and older engaged in heavy drinking or other drug use¹⁴
- Percentage of women who gave birth in the past year and reported using alcohol or other drugs during pregnancy
- Mean age at first use of “gateway” drugs (tobacco, marijuana, alcohol)
- Percentage of adolescents aged 14-17 stating disapproval of marijuana use
- Percentage of adolescents aged 14-17 who report parents or guardians who communicate non-use expectations
- Percentage of drug abuse clients who engage in risk behaviors related to HIV/AIDS after completing treatment plan

Sexual Assault Prevention

- Incidence rate of sexual assault reported by females

Disabilities

- Percentage of newborns with neural tube defects
- Percentage of persons aged 18-65 with disabilities who are in the workforce
- Percentage of children aged 6 or younger with blood lead greater than 10 micrograms per deciliter¹⁵
- Percentage of women who gave birth in the past year and reported using alcohol, tobacco, or other drugs during pregnancy

Emergency Medical Services

- Percentage of persons who suffer out-of-hospital cardiac arrest who survive

¹²Although the estimated incidence rate would be a more appropriate measure for monitoring progress by the state substance abuse agencies, the currently available data source for this measure provides prevalence data.

¹³See fn. 12.

¹⁴See fn. 12.

¹⁵See fn. 1.

APPENDIX B

Workshop Agenda

July 22, 1997

National Academy of Sciences
The Lecture Room
2101 Constitution Avenue, N.W.
Washington, DC

8:30 a.m.	Welcoming Remarks and Introductions	Edward Perrin Susanne Stoiber
-----------	-------------------------------------	----------------------------------

Presentations from Invited Guests on Data Systems

9:00	Department of Health and Human Services (DHHS)	Dale Hitchcock
------	--	----------------

9:30	National Center for Health Statistics (NCHS)	Edward Sondik
------	--	---------------

10:00	Centers for Disease Control and Prevention (CDC)	Gary Hogelin Chuck Gollmar
-------	--	-------------------------------

10:30	<i>Break</i>	
-------	--------------	--

10:45	National Committee on Vital and Health Statistics, DHHS	Don Detmer
-------	---	------------

11:15	General Discussion	Edward Perrin
-------	--------------------	---------------

12:00 p.m.	<i>Working Lunch (Discussion of Federal-State Data Systems Issues: Examples from Education)</i>	John Ralph William Morrill
1:30	Substance Abuse and Mental Health Services Administration (SAMHSA)	Winnie Mitchell Peggy Gilliam Samuel Korper Charleen Lewis
2:00	National Association of State Mental Health Program Directors (NASMHPD)	Vijay Ganju
2:30	University of Pennsylvania and TRI-Net	Thomas McLellan
3:00	<i>Break</i>	
	<i>Continuation of Presentations from Invited Guests on Data Systems</i>	
3:15	Science and Epidemiology Committee, Association of State and Territorial Chronic Disease Program Directors	Christopher Maylahn
3:45	Health Care Financing Administration Overview of Medicaid and Medicare Information Systems and Surveys	Michael Hadad
4:30	General Discussion	Edward Perrin
5:00	Adjournment	

APPENDIX C

Biographical Sketches

EDWARD B. PERRIN (*Chair*), a biostatistician and health services researcher, is professor and former chair of the Department of Health Services and of the Department of Biostatistics at the University of Washington, and senior scientist at the V.A. Puget Sound Health Care System. He is a member of the Institute of Medicine and of the Committee on National Statistics of the National Academy of Sciences. He is a former director of the National Center for Health Statistics and past president, member of the Governing Board, and distinguished fellow of the Association for Health Services Research. Dr. Perrin has served as chair of the Health Services Research Study Section and as a member of the National Advisory Council for the Agency for Health Care Policy and Research, DHHS, and is currently chair of the Scientific Advisory Committee at the Medical Outcomes Trust. His research and teaching interests and scientific publications focus on the development of new methodologies for the measurement of health outcomes and the use of large health data systems in decision making and policy development. He received a B.A. in mathematics from Middlebury College, an M.A. in mathematical statistics from Columbia University, and a Ph.D. in statistics from Stanford University.

JOHN C. BAILAR III is professor and former chair of the Department of Health Studies, University of Chicago. In earlier years he was at the National Cancer Institute, Harvard University, and McGill University. He is trained in both medicine (M.D.) and statistics (Ph.D.), and works mainly on problems in the intersection of these fields. His special interests include cancer epidemiology, data quality, scientific communication, and misconduct in science.

THOMAS J. BARRETT has been director/commissioner for Colorado Mental Health Services since February 1996. He is responsible for planning, policy implementation, and resource allocation for the Colorado public mental health system, including community service contracts for over \$150 million. He is also an adjunct full professor at the University of Denver and a member of the clinical faculty at the University of Colorado Health Sciences Center. He is currently an officer for the National Association of State Mental Health Program Directors and cochairs that organization's President's Task Force on Performance Measures. He holds a Ph.D. in clinical psychology and has published numerous articles in the areas of mental health administration, performance evaluation, and program evaluation.

JANE S. DURCH joined the staff of the Committee on National Statistics in November 1997. Previously she was at the Institute of Medicine (IOM), where she participated in several studies on subjects including community health and performance monitoring, vaccine development, and emergency medical services for children. Before coming to the IOM, she worked with the Division of Health Policy, Research, and Education at Harvard University; the Committee on Population and Demography at the National Research Council; and the Population Reference Bureau. She has a B.S.F.S. from the Georgetown University School of Foreign Service and an M.A. in demography from Georgetown University.

DAVID W. FLEMING is Oregon's state epidemiologist and oversees Health Division programs in communicable disease prevention; HIV and sexually transmitted diseases; health promotion and chronic disease prevention; environmental, occupational, and injury epidemiology; and vital records and community assessment. He received his M.D. at the State University of New York, his training in internal medicine at Oregon Health Sciences University (OHSU), and his training in preventive medicine at the Centers for Disease Control and Prevention. He is on the faculty of the Department of Public Health and Preventive Medicine at OHSU and has authored or coauthored a number of scientific publications. He sits on several national advisory boards and task forces, including the Advisory Committee on Immunization Practices and the U.S. Public Health Service Task Force on Community Preventive Services. He is the immediate past president of the Council of State and Territorial Epidemiologists, the national association of state epidemiologists.

V. JOSEPH HOTZ is a professor in the Departments of Economics and Policy Studies at UCLA. Prior to his current appointment, he was a founding faculty member of the Irving B. Harris Graduate School of Public Policy Studies at the University of Chicago. While at Chicago, he served as a research associate and director of the Population Research Center. He was involved in the formation of the Northwestern/University of Chicago Joint Center for Poverty Research, for

which he is now a national research associate and chair of the Advisory Panel for Research Uses of Administrative Data. He also serves as a research associate of the National Bureau of Economic Research, and has served on several National Academy of Sciences panels and as a member of the U.S. Secretary of Health and Human Services' JOBS Evaluation Advisory Panel. Over the past 5 years, he has had an ongoing advisory role on welfare reform with the California Department of Social Services, serving on its Research Advisory Committee and, more recently, on its Welfare Reform Data Collection Advisory Group. He holds a Ph.D. in economics and has published extensively in the areas of the economics of the family, applied econometrics, and the evaluation of social programs.

IRA R. KAUFMAN is clinical associate professor at the University of Medicine and Dentistry of New Jersey, and he teaches at the New Jersey School of Public Health. Currently, he is the director of the W. K. Kellogg Foundation Information Access Program, and he was formerly director of the Robert Wood Johnson Foundation Information for State Health Policy Program. He has advised various federal agencies, including the Centers for Disease Control and Prevention, the Agency for Health Care Policy and Research, the Health Resources and Services Administration, and the Health Care Financing Administration, as well as the Deputy Assistant Secretary for Health Policy. He is a member of the advisory board of the American Hospital Association's Committee on Health Statistics and the executive board of the National Association of Health Data Organizations. He also served on the National Information Infrastructure Health Care Advisory Group of the Council on Competitiveness. Previously, he held positions in the Wisconsin Department of Health and Social Services, where he directed research and data activities for statewide health data systems. Also, he served as a primary consultant on data issues for Wisconsin's Medicaid program, public health programs, and regulatory programs. He has a B.S. and M.S. in sociology from the University of Wisconsin, Madison.

JOHN R. LUMPKIN was appointed director of the Illinois Department of Public Health in January 1991. He is the first African-American to hold this position. Previously, he was associate director of the department's Office of Health Care Regulation. Before joining the state health department, he was an emergency physician at several Chicago hospitals. He is an expert on public health information systems and performance measurement and teaches these subjects at the University of Illinois at Chicago, School of Public Health. He is also a leading expert on injury prevention. He has served on several national advisory committees and currently serves on the National Committee on Vital and Health Statistics and the Centers for Disease Control and Prevention's Advisory Committee to the Director. He was a member of the Institute of Medicine Committee on Using Performance Monitoring to Improve Community Health and the Public Health Roundtable. He is past president of the Association of State and Territorial Health

Officials and a current member of the executive committee, a member of the board of trustees of the Foundation for Accountability, a past board member of the American College of Emergency Physicians, and past president of the Society of Teachers of Emergency Medicine. He received his M.D. in 1974 from Northwestern University Medical School, where he continues to serve as assistant professor in emergency medicine. He trained in emergency medicine at the University of Chicago and earned his M.P.H. from the University of Illinois at Chicago, School of Public Health.

WILLIAM A. MORRILL is a senior fellow at Mathtech, Inc., a social science applied research consulting firm. As a former Assistant Secretary for Planning and Evaluation at the U.S. Department of Health, Education, and Welfare, he developed a thorough grounding in the nation's public health programs during his 4 years of service. He served more than a decade at the Office of Management and Budget, and also as deputy county executive of Fairfax County, Virginia. His applied research work has been concentrated in education, health, and human services; many of his recent projects have focused on performance measurement, strategic planning, and evaluation. He is a member of the National Academy of Public Administration (NAPA) and has served in leadership roles in NAPA and the Council for Excellence in Government. He has also served on many commissions, standing committees, and panels of the National Research Council. He holds a B.A. from Wesleyan University and an M.P.A. from the Maxwell School, Syracuse University.

R. HEATHER PALMER is director of the Center for Quality of Care Research and Education in the Department of Health Policy and Management at the Harvard School of Public Health. She led the research team funded by the Agency for Health Care Policy and Research (AHCPR) that built the prototype system for the collection, retrieval, and evaluation of clinical performance measure sets, which is now available through the AHCPR home page as CONQUEST, the Computerized Needs-oriented Quality Measurement Evaluation SysTem. CONQUEST is being developed as a national resource through the Quality Measurement Network (QMNet) project, funded by AHCPR. Dr. Palmer's current research focuses on measuring and improving quality of care. She serves on the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) Advisory Council on Performance Measurement, for which she is cochair of the Measure Evaluation Subcommittee; the Performance Measurement Advisory Committee of the American Medical Accreditation Program of the American Medical Association (AMA); and the Performance Measurement Coordinating Council established by the AMA, JCAHO, and the National Committee for Quality Assurance. Dr. Palmer is editor-in-chief of the journal of the International Society for Quality in Health Care. She earned her baccalaureate from Cambridge University, the M.B. and B.Ch. (equivalent to the M.D. in the United

States) from Cambridge University and the Royal London Hospital Medical College, and an M.S. in health services administration from the Harvard School of Public Health.

J. SANFORD SCHWARTZ is professor of medicine and health management and economics at the School of Medicine and the Wharton School, University of Pennsylvania. He is also a senior fellow and former executive director of the Leonard Davis Institute of Health Economics. He received an A.B. in history from the University of Rochester and an M.D. from the University of Pennsylvania, and completed the Wharton School's M.B.A. program following a residency in internal medicine at the University of Pennsylvania. His research focuses on the evaluation of medical practices and medical decision making, including evaluating trade-offs among cost, quality, and outcomes in health care and optimizing the value of clinical information. He has received research support from and served as an advisor or consultant to a broad spectrum of governmental and private-sector groups, including the Centers for Disease Control and Prevention; the Department of Defense; the Health Care Financing Administration; the Institute of Medicine; the National Institutes of Health; the U.S. Congress Office of Technology Assessment; the U.S. Preventive Services Task Force; the Veterans Administration; the World Health Organization; health insurers and managed care organizations; the White House Health Care Reform Task Force; and many foundations, private-sector groups, and corporations. He is past president and trustee of the American Federation for Clinical Research and the Society for Medical Decision Making. He currently serves as a trustee of the American Society for Hypertension, associate editor of the *Journal of General Internal Medicine*, and editor-in-chief of the *American Journal of Managed Care*.

SUSAN M. SKILLMAN is administrator of the Center for Cost and Outcomes Research at the University of Washington. She has been with the university for 9 years and has directed a wide range of research programs and projects, including the university's Center for Rural Health Services Research, studies of risk adjustment for managed care capitation, and studies of access to health insurance for low-income populations. Previously she worked for 8 years at the Center for Health Studies at Group Health Cooperative of Puget Sound, where she was coordinator for research development and managed several research projects. She has a B.A. from Whitman College and an M.S. from Washington State University.

MARY E. STUART is executive director of the Maryland Collaboratory for Business and Health. She is an associate professor in the Department of Sociology and Anthropology at the University of Maryland, Baltimore County, where she teaches courses on the organization and delivery of health services and issues in health care. She also holds faculty appointments at the Johns Hopkins School of Hygiene and Public Health and the University of Maryland School of Phar-

macy. Previously, she was director of policy for the Maryland state health department and director of Medicaid policy. She has been involved in numerous research studies of factors affecting the cost and outcomes of health care. These include recently completed studies to examine factors contributing to high costs for Medicaid patients, a randomized clinical trial of outreach strategies for pregnant substance abusers, and patient satisfaction studies for Medicaid. She was active in promoting passage of Maryland's health reform legislation in 1993, and was appointed by the Governor to chair the Data Base Work Group of the Maryland Health Care Access and Cost Commission. She has a Doctor of Science degree in health policy and management from the Johns Hopkins School of Hygiene and Public Health.

CYNTHIA P. TURNURE is health care program manager for the Center for Health Statistics, Minnesota Department of Health. Previously, she served as director of Minnesota's Fetal Alcohol Syndrome Coordinating Board and as executive director of the Chemical Dependency Program Division of the Minnesota Department of Human Services, a position she held for 15 years. Under her leadership, Minnesota developed a number of state-of-the-art information systems that are considered models for the nation. She also helped develop Minnesota's Consolidated Chemical Dependency Treatment Fund, which combined a number of previously separate state and federal funding sources into one fund that allowed the "dollar to follow the client" based on uniform assessment and placement criteria. She served as a member of the Institute of Medicine committee that produced the 1990 report *Broadening the Base of Treatment for Alcohol Problems*. She has a B.A. from Wellesley College and a Ph.D. in developmental psychology from Yale.

PAUL J. WIESNER is director of the DeKalb County (Georgia) Board of Health. He completed his medical training at the University of Wisconsin and his internal medicine and infectious disease training at the University of Rochester and the University of Washington. During a 20-year career with the Centers for Disease Control and Prevention (CDC), he directed national programs on the control of sexually transmitted diseases, chronic disease, and public health training. He is on the board of the National Association of County and City Health Officials (NACCHO) and the Center for the Advancement of Community Based Public Health. He is a member of the faculty of the Rollins School of Public Health of Emory University, the CDC Preventive Medicine Advisory Council, and the Coordinating Council of the Coalition for Healthier Cities and Communities. He is also chair of the Council of the Public Health Leadership Society. As chair of the APEX Workgroup for NACCHO, he is committed to developing local community systems for public health and providing practical tools for the transformation of local health departments. He is the author of more than 70 articles in scientific journals.

GAIL R. WILENSKY chairs the Medicare Payment Advisory Commission and serves as the John M. Olin Senior Fellow at Project HOPE, where she analyzes and develops policies relating to health reform and to ongoing changes in the medical marketplace. In both capacities, she testifies frequently before congressional committees; acts as an advisor to members of Congress and other elected officials; and speaks nationally and internationally before professional, business, and consumer groups. Previously, she served as deputy assistant to President Bush for policy development, advising him on health and welfare issues. Prior to that, she was administrator of the Health Care Financing Administration, overseeing the Medicare and Medicaid programs. She is a member of the Institute of Medicine and serves as a trustee of the Combined Benefits Fund of the United Mineworkers of America and of the Research Triangle Institute. She is an advisor to the Robert Wood Johnson Foundation and the Commonwealth Fund and is a director on several corporate boards. She received a bachelor's degree in psychology and a Ph.D. in economics at the University of Michigan.