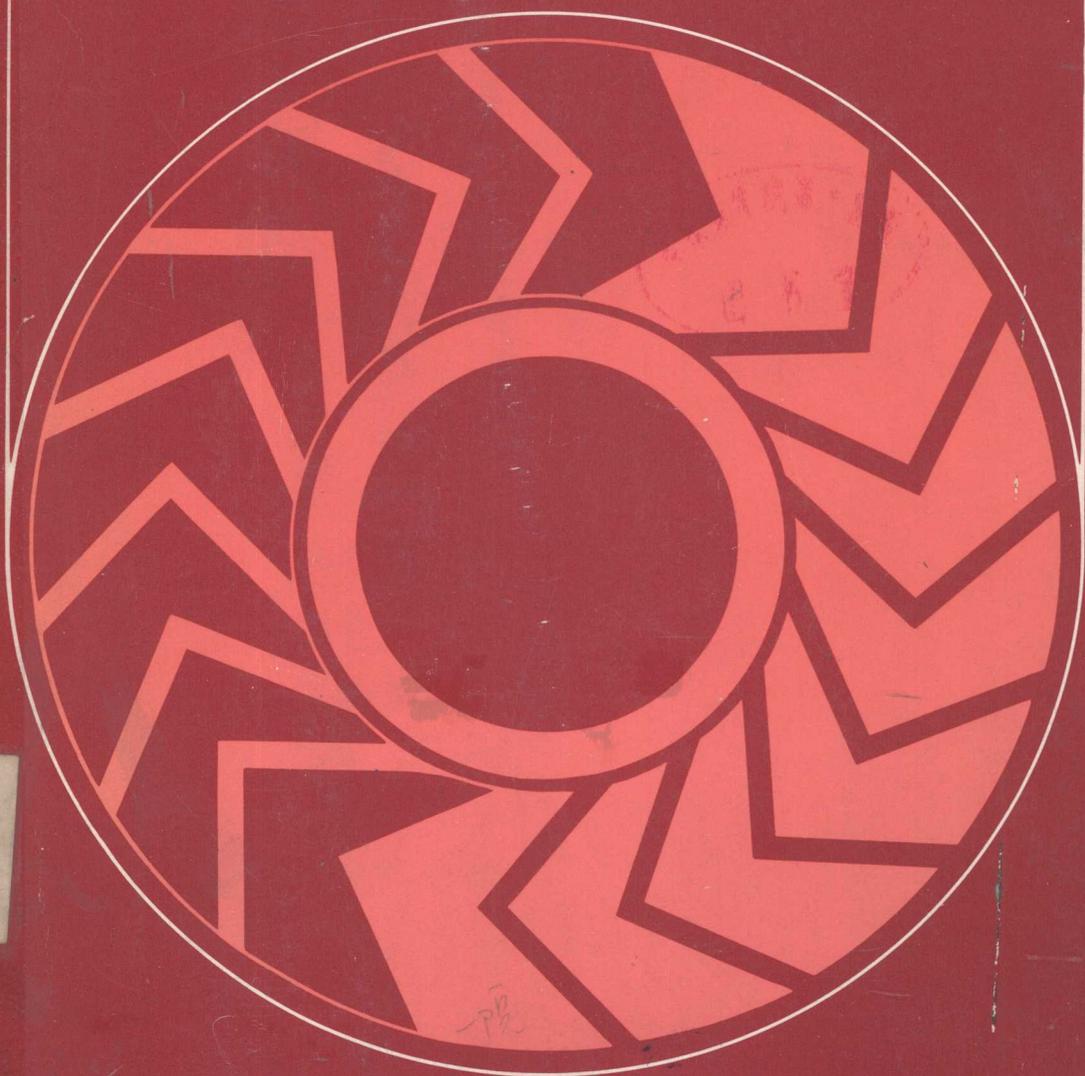


# Health program evaluation

Stephen M. Shortell  
William C. Richardson



# Health program evaluation

**Stephen M. Shortell, Ph.D.**

**William C. Richardson, Ph.D.**

Department of Health Services,  
School of Public Health and Community Medicine,  
University of Washington, Seattle, Washington

*with 14 illustrations*



**The C. V. Mosby Company**

Saint Louis 1978

Copyright © 1978 by The C. V. Mosby Company

All rights reserved. No part of this book may be reproduced in any manner without written permission of the publisher.

Printed in the United States of America

The C. V. Mosby Company  
11830 Westline Industrial Drive, St. Louis, Missouri 63141

**Library of Congress Cataloging in Publication Data**

Shortell, Stephen M.

Health program evaluation.

(Issues and problems in health care)

Bibliography: p.

1. Health services administration—Evaluation.
  2. Public health—Evaluation. 3. Evaluation research (Social action program) I. Richardson, William C., joint author. II. Title. [DNLM: 1. Health services—United States. 2. Public health administration—United States. 3. Evaluation studies. W84 AA1 S513h] RA394.S56 362.1'04'25 78-4866
- ISBN 0-8016-4595-6

C/M/M 9 8 7 6 5 4 3 2 1

# Health program evaluation



## **Issues and problems in health care**

**Paul R. Torrens, M.D., M.P.H., Series editor**

School of Public Health,  
University of California,  
Los Angeles, California

### **Other available titles**

*Rural Health Care*, Milton I. Roemer, M.D.

*Elements of Planning for Area-Wide Personal Health Services*,  
William Shonick, Ph.D.

*The Law and the Public's Health*, Kenneth R. Wing, J.D., M.P.H.

*National Health Policy and the Underserved: Ethnic Minorities, Women,  
and the Elderly*, Jerry L. Weaver, Ph.D.

*The American Health Care System: Issues and Problems*,  
Paul R. Torrens, M.D., M.P.H.

### **Forthcoming title**

*Issues in Women's Health Care*, Helen I. Marieskind, Dr. P.H.

*For*  
**Susan and Nancy**  
*and our children*  
Stephanie, Juliana, Elizabeth, and Jennifer



## Foreword

During the past decade, the cost for health care has been escalating at an alarming rate. In 1976, health care expenditures in the United States reached \$139.3 billions, or 8.6 percent of the gross national product, compared to \$4.2 billions or 5.8 percent ten years ago. As expenditures increased, the public sector's share rose from 25.7 percent of total expenditures in 1966 to 42.2 percent in 1976. This rapid increase has caused federal, state, and local officials to critically examine current and planned health care programs in order to make careful decisions about priorities and allocation of limited resources. Program evaluation is important to policymakers because it provides the sound factual basis for making these decisions. Whether we are determining how hospitals should be reimbursed or alternatives to institutional long-term care, program evaluation can provide a fresh, objective examination of the problems and issues involved. It can identify which programs are performing well or badly and why. Furthermore, program evaluation can furnish the data and information needed to strengthen weak programs, support effective programs, or eliminate programs that are not meeting legislative and/or agency goals and objectives. This feedback is extremely important to legislators and government officials.

Professors Stephen M. Shortell and William C. Richardson are to be commended for their interest in equipping present and future health service professionals for the task of dealing with questions of health program design, analysis, and evaluation. Their text underscores the compelling need for researchers to understand the political and administrative environment in which research and evaluation dealing with health service programs are likely to occur, since the nature of that milieu can contribute to the shape of the evaluation outcome. It attempts to bridge the two diverse worlds of the basic technical evaluation process on the one hand and, on the other, the political and administrative realities within which the assessment is conducted.

The authors have, to their credit, distinguished among various types of program settings and organizational issues that confound the evaluation process and that too often limit results despite the best design preparation and intention.

The Health Care Financing Administration (HCFA), one of the most recently created agencies of the Department of Health, Education, and Welfare, is especially concerned with effective health program evaluation through its Office of Policy, Planning, and Research. Since it assumed the combined responsibility for the administration of the Medicare and Medicaid programs,

for professional standards review, and for research and statistics related to health care financing, its concerns have extended beyond the funding and management of a massive share of the nation's medical care services. The cost of health care, and its continuing escalation, impelled in part by new federal programs, requires that we attempt to test promising new proposals for improved financing, delivery, and organization of medical care before they reach the stage of legislative or operational readiness. Our ongoing experience must also be carefully examined for highlights leading to more effective management and public policy.

An increasing awareness of the evaluative function, however, pervades the entire health services structure so that administrators, program planners, and providers must be more directly concerned with the consequences of their actions as they seek to change health care organizations and improve medical care practices. Accountability by public officials as well as medical care providers is the inevitable consequence of growing consumer sophistication and public financing. Shortell and Richardson are keenly aware of this new responsibility that is being thrust upon us, and this text should prove a valuable resource for both current and future administrators, providers, and policymakers in meeting the challenge.

**Robert A. Derzon**

*Administrator*

**Clifton R. Gaus**

*Associate Administrator for  
Policy, Planning, and Research*

Health Care Financing Administration,  
Department of Health, Education, and Welfare

## Foreword

For many of us, it is impossible to give up completely the stereotypes we hold about medical research and health care. The great medical discoveries of the 1800's, the magical control of tropical diseases during that century and the early decades of this one, and the prevention of communicable diseases by community-wide vaccination programs were truly remarkable demonstrations of human persistence and creativity. These events, reinforced by sometimes factually based and othertimes scientifically unsupported popular novels, and of course by the glamorization of practicing physicians and by the dramatization of the miracles of modern medical care on television, are responsible for the popular views on the power of medical research and the efficacy of health care.

It is important, however, that the stereotypes of medical research and health care not be accepted unquestioningly. As social medicine experts and medical historians have informed us repeatedly, most of the decisive reductions in mortality and morbidity are accounted for by improvements in general living conditions; moreover, as reviews of accomplishments in disease control report, new medical interventions and technological advances typically result in only marginal improvements in health status. Certainly now, except for a few disease areas and in some underdeveloped countries, reduction of morbidity and mortality is an incremental, not a revolutionary, process. Further, increasingly, the chief architects of improved health status and the important warriors against illness and disease are the specialists identified with community medicine, public health and the delivery of health services.

Surely, there will be additional laboratory achievements in medicine and more technological breakthroughs in diagnosis and treatment. But it is a safe prediction that future improvements in the health and comfort of community members, even more than in the past, will be accomplished by attacking the defective conditions of our social milieu and physical environment, by remedying the personal and interpersonal behaviors of community members so that prevention and compliance increase, and by reorganizing the delivery of health services so that access to and continuity of competent health care is universally available. It is these areas that are emphasized in current health initiatives to expand and restructure inner city ambulatory care and increase the number of competence of rural health providers, in the emergence of a variety of new health practitioners, in the revitalization of health education, in attacks on industrial and environmental pollution, in the expansion of mental health activities, and in the press for expanded governmental support of the costs of health care.

In some areas, the efforts of the last decade are impressive, as in the markedly increased access to care of the urban poor and minority group members. In others it is disheartening, as in the cases of tobacco smoking, drug addition, obesity, and many other health-threatening behaviors. Moreover, some of the efforts that have proved successful in terms of their objectives have been accompanied by fearsome side effects, including the spiraling of costs of health services and consumer disenchantment with the health care system. Economic resource constraints as well as the power of social traditions and cultural norms, not to mention the rigidities of extant bureaucratic arrangements and the self-interests of various parties within the health industry, impede efforts directed at the goal of improved health status for community members. The dictum that gains in the health field usually are incremental and modest and that many initiatives and innovative programs are, on balance, either ineffective or inefficient, must be accepted now as in the past, despite our romantic stereotypes of progress in medical care and health status.

Rather than to "eyeball" the effects of programs and initiatives in public health and medical care, and to depend on judgmental impressions of the utility of different delivery of health services approaches, it is essential that as systematic, replicable, and precise assessments as possible be undertaken of both established and innovative efforts. Unless the press is toward such assessments, modest gains may be overlooked and either ineffective or overly costly interventions adopted, with consequent loss of support by disgruntled influentials and political and community groups.

Besides, data, evidence, and the rigor of scientific inquiry have characterized the adoption of new procedures and practices in the health area since the beginnings of modern medicine. That "proof" needs to be obtained by a set of rules is not disagreeable either to the laboratory scientist or the clinical investigator, and it should not be for the health professional engaged in the innovation and conduct of macro-activities. This is all evaluation research is, a set of procedures that when properly implemented allow inference, with as much certainty as possible, to be drawn about the feasibility, efficacy, and efficiency of various approaches to the prevention, control, and management of disease and the organization and arrangement for the delivery of health services.

Evaluation research, in a sense, is an organizing concept. In its basic perspective, it makes use, as much of this text indicates, in the underlying commitment to the "experiment," to the basic outlook of the laboratory scientist. In its data collection approach, as other parts document, its procedures often are the same as those employed by the epidemiologist. The practicing professional, the health science student, and the medical investigator, then, will feel comfortable with much of the material included here. The chapters, in this perspective, represent a codification and a restatement of how to accumulate evidence, how to make inferences and how to utilize research for health care interventions.

At the same time, the volume contains much new material for the health professional and researcher, for evaluation research is rooted to a large extent in social research methodologies. Moreover, evaluation research, as an out-

look and a set of procedures, is equally commonplace in welfare, education, and public safety efforts—indeed, across the entire human services arena—as in the health field. Many of the ideas, principles, and techniques presented are drawn from the accumulating body of evaluation research being undertaken in different fields. The authors have strived, however, to create a text about health evaluations and have employed and adopted examples and illustrations to meet the objective of providing a special book for the field.

To my knowledge, this is the first basic text on *health* evaluations. Not only does its specialized character make instruction in evaluation research more relevant to persons in the health field, but it brings together for the professional and investigator procedures and viewpoints previously scattered in the elusive and varied evaluation research literature. The authors have also strived to make the book short without sacrificing completeness, and readable without being over-simple.

One final observation: the successful conduct of evaluation research requires attention to practical as well as technical details. Evaluations are exciting but difficult to do well because of the contingencies that surround doing research in the complex community and organizational world in which health care activities are lodged. Many of the pages of this text provide advice in these practical matters—the section on how to state program objectives is but one example. Readers of this text, in undertaking their own evaluations, and in appreciating the results of work done by others, I am certain, will come to appreciate the effort to consider both the technical and pragmatic elements involved in evaluation research.

**Howard E. Freeman**

Institute for Social Science Research,  
University of California,  
Los Angeles, Calif.

## Preface

Over the past several years, we have been teaching program evaluation to graduate students interested in the administration and delivery of health services. A major difficulty in teaching this material involves the wide variation in students' academic backgrounds, resulting in problems of what to include and exclude and at what level to present the material. The problem is complicated by an evaluation literature scattered throughout widely different sources. For a discussion of experimental and quasi-experimental designs, one draws on Campbell and Stanley (*Experimental and Quasi-Experimental Designs for Research*); for discussion of measurement reliability and validity issues, the educational psychology literature is helpful; for discussion of evaluation design in health services, Suchman (*Evaluation Research: Principles and Practice in Public Service and Social Action Programs*) is a useful source; for empirical examples, several collections of readings are available (for example, Caro, *Readings in Evaluation Research*; Schulberg and others, *Program Evaluation in the Health Fields*; Mullen, *Evaluation of Social Intervention*; and Struening and Guttentag, *Handbook of Evaluation Research, Volume II*). One goes to still other sources for discussion of program implementation (for example, Williams and Elmore, *Social Program Implementation*), the administrative and political issues surrounding program evaluation (for example, Weiss, *Evaluation Research*), and the role of program evaluation in the development and implementation of public policy.

To a great extent this diversity reflects an increased interest in program evaluation itself, and it should be applauded. It is up to the instructor to integrate the material in a way that contributes to the learning objectives of the students. But we believe the process can be expedited by developing course material that brings together "in one place" some of the key concepts, methodologies, and issues related to program evaluation in general and their applications to the delivery of health services in particular.

The major purpose of this textbook, then, is to provide a systematic presentation of the major concepts, methodologies, and issues concerning the evaluation of health services delivery programs at a level appropriate to graduate students in health services administration and planning, public health, dentistry, medicine, nursing, pharmacy, social work, and related health professions. As Henderson and Meinert note: "Most of the textbooks in the fields of both epidemiology and biostatistics devote little space to design and analysis questions related to health and medical evaluation." This book is thus aimed primarily at individuals who will eventually be involved in the administra-

tion, planning, delivery, and evaluation of health services programs. The text will also be a useful first source to discipline-oriented students (for example, in psychology, sociology, education) or others who wish to specialize and become "experts" in evaluation research itself. For such students, however, the material in this text would need to be supplemented by additional readings and coursework.

This book also attempts to redress an imbalance caused by the somewhat artificial distinction between technical/methodological evaluation issues and the political and administrative settings in which evaluation is conducted. It is our belief that the future administrator, planner, or provider of health services needs to know something about both evaluation research itself and the environment in which it takes place. For us, the relevant question is not should future administrators, planners, and providers be trained to do research, but rather, what does an understanding of the evaluation research process contribute to their overall competence and effectiveness? It is hoped that the sensitivity of the health science student to the importance of program evaluation will be increased, together with a further understanding of the uses and limitations of program evaluation in the decision-making and policy-making processes.

The contents of the book are organized in a sequence that we have found useful in communicating with students. However, the individual chapters are sufficiently self-contained that other instructors can use the material in a different order or can assign specific chapters at different points in the course. The first chapter traces some of the early background and historical development of program evaluation efforts in the health care field and describes some contemporary forces influencing the current shape and content of health program evaluations. In Chapter 2, the student is introduced to issues concerning the development of program objectives, different levels and types of objectives, and the design of program components. In Chapter 3, some major experimental and quasi-experimental designs applicable to health program evaluation are discussed. Emphasis is placed on the pros and cons of each design relevant to issues of internal and external validity. Chapter 4 then builds on Chapter 3 by discussing issues related to the reliability and validity of individual measures, the advantages and disadvantages of different methods of data collection, and basic data analysis approaches. In Chapter 5, the important administrative and political issues in program evaluation are discussed, along with an analysis of the problems of program evaluation. Chapter 6 concludes with a discussion of future issues in program evaluation, with an emphasis on the role of program evaluation in developing and implementing public policy in the delivery of health services.

At the end of each chapter, where pertinent, a glossary of terms is presented, along with class problem exercises and a list of suggested readings. In addition an Appendix presents an example of a student's evaluation research paper. We have found the problem exercises, together with the requirement that a student develop an evaluation of an ongoing delivery program, to be valuable learning experiences that reinforce the course readings and classroom discussions.

From the preceding comments and, hopefully, from the material to follow, it may appear that this book represents a rational approach to the subject. However, it was not written in a particularly rational environment. Contrary to what may prevail elsewhere, and certainly contrary to common opinion, this book was not written in the quiet solitude of an academic leave of absence, sabbatical, or even an “off quarter” from teaching, research, or administrative responsibilities. Rather, it was written in hotel rooms, on airplanes, in airports, on buses, between classes and student counseling sessions, between faculty meetings, and, yes, *in* faculty meetings! Thus, whatever clarity and coherence may exist in the chapters that follow are due in small part to our perseverance and in large part to the skill and careful attention of our editor, Augie Podolinsky, the secretarial support provided by Bernice Goldberg and Elaine Morrissey, and the library research provided by Diane McKenzie. The book has also benefited greatly from the comments and suggestions of our colleagues at the University of Washington—Allan Blackman, Marilyn Bergner, and Walt Williams—and, in particular, from the advice of Ron Andersen, University of Chicago, and Don Riedel, presently University of Washington and formerly Yale University. It is entirely possible that a key point raised by them has been lost somewhere along the way (perhaps at an airport terminal), but as teachers of program planning and evaluation in their own right, they will be able to remedy our neglect or shortcomings. This, of course, applies to all future users of this text.

Finally, as implied already, this text represents a beginning. Since the ultimate test of program evaluation itself is the extent to which it is actually used by, and is useful to, those who have to make program decisions, so too the merit of this text must be judged by those who find it useful in teaching program evaluation to health science students. The ultimate impact, of course, must be judged by the increasing numbers of health science students (administrators, planners, direct providers) who have an understanding and a working knowledge of program evaluation methodologies and who can carry this over into their professional practice. We are, of course, not so optimistic or naive to think that this in itself will have an impact on the health status of the United States population. Rather, we hope it may help everyone (administrator, planner, provider, and consumer) to better understand the uses and limits of health service in a complex society.

**Stephen M. Shortell**

**William C. Richardson**

Someone has suggested that evaluation is like salt.  
Some foods need more salt than others,  
but too much salt can ruin a meal.

**A. L. Knudson,**

*"Evaluation for What?"*

Dogma is the enemy of truth and the enemy of persons.  
The ideas enshrined in dogma may include good and wise ideas,  
but dogma is bad in itself because it is accepted  
as good without examination.

**OK Words**

# Contents

- 1 Program evaluation: historical antecedents and contemporary developments, 1
  - 2 The evaluation process, 16
  - 3 Evaluation designs, 38
  - 4 Measurement, data collection, and data analysis issues, 74
  - 5 Program evaluation and problems of implementation, 98
  - 6 Future issues: evaluation research and public policy, 121
- APPENDIX** Multiple time series evaluation of a cervical cancer screening program, 134  
**Sharon Baker**

## CHAPTER 1 Program evaluation: historical antecedents and contemporary developments

The purpose of this chapter is to introduce students and practitioners in the health sciences to some basic developments in the evolution of program evaluation. A *program* is defined here as “. . . an organized response to eliminate or reduce one or more problems where the response includes one or more objectives, performance of one or more activities, and expenditure of resources.”<sup>1</sup>

A number of terms will be defined and compared. These include (1) evaluation research, (2) nonevaluative research, (3) policy research, (4) policy analysis, (5) impact or “summative” evaluation, and (6) process or “formative” evaluation. Factors shaping both the growth and nature of future program evaluation will also be explored. The many motivations for conducting program evaluation activities will be discussed. Finally, the relevance of program evaluation for practicing administrators, planners, and providers will be noted. Comprehension of the material in this chapter provides a framework for understanding future chapters and will enhance the reader’s eventual ability to analyze, apply, synthesize, and evaluate what is to be learned from this text.

### Historical development

Attempting to evaluate a social program is at best risky and at worst treacherous. In few other activities are the ambivalent tendencies of society so clearly revealed. On the one hand is society’s desire to learn more so that the quality of life may be improved, while on the other hand is the ubiquitous fear of what might be found. It is a phenomenon similar to individual growth and development, but it is acted out and institutionalized at the level of social collectives with consequences that frequently have far-reaching implications for large numbers of people. This tug-of-war between the desire to learn more and the need for self-protection is frequently the genesis of social conflict and social change, and is clearly revealed as one traces the evolution of health services over the past several centuries. A brief review of this evolution is important *not* because history might repeat itself, but because what is learned from previous experiences, with adaptation, may be applicable in present circumstances.

Evaluation of medical care services has existed in some form from earliest times. We know that such evaluations were often closely tied to sanctions. For