EVALUATING THE HEALTHCARE SYSTEM:

Effectiveness, Efficiency, and Equity

Third Edition

Lu Ann Aday
Charles E. Begley
David R. Lairson
Rajesh Balkrishnan

Health Administration Press, Chicago, Illinois
AcademyHealth, Washington, DC

Your board, staff, or clients may also benefit from this book's insight. For more information on quantity discounts, contact the Health Administration Press Marketing Manager at (312) 424-9470.

This publication is intended to provide accurate and authoritative information in regard to the subject matter covered. It is sold, or otherwise provided, with the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

The statements and opinions contained in this book are strictly those of the author and do not represent the official positions of the American College of Healthcare Executives or of the Foundation of the American College of Healthcare Executives.

Copyright © 2004 by the Foundation of the American College of Healthcare Executives. Printed in the United States of America. All rights reserved. This book or parts thereof may not be reproduced in any form without written permission of the publisher.

08 07 06 05 04 3 2 1

Library of Congress Cataloging-in-Publication Data

Evaluating the healthcare system: effectiveness, efficiency, and equity / Lu Ann Aday . . . [et al].—3rd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 1-56793-222-3 (alk. paper)
RA399.A5E95 2004 362.1 068—dc22

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.®TM

Acquisitions manager: Audrey Kaufman; project manager: Joyce Sherman; cover design: Megan Avery

Health Administration Press
A division of the Foundation of the American College of Healthcare Executives
One North Franklin Street
Suite 1700
Chicago, IL 60606
(312) 424-2800
Chapter 2

Effectiveness: Concepts and Methods

CHAPTER HIGHLIGHTS

1. Effectiveness is the degree to which improvements in health now attainable are, in fact, attained. Effectiveness concerns the results achieved in the actual practice of healthcare with typical patients and providers, in contrast to efficacy, which is assessed by the benefits achieved under ideal conditions.

2. Effectiveness research reflects two seemingly competing, but in fact complementary, definitions of effectiveness. One represents a population perspective, and the other represents a clinical perspective. The four levels of measuring effectiveness include the community level, associated with the population perspective, and the system, institution, and patient levels, associated with the clinical perspective.

3. Structure, process, and outcomes are linked conceptually in a research paradigm that assumes structural elements of healthcare as having an influence on what is and is not done in the process as well as how well it is done; this process in turn influences the health outcomes people experience as a result of their encounters with the care delivery process.

4. Outcome measures fit into the general categories of mortality, morbidity, and health status, although specific indicators of each of these broad measures may be used at different levels.

5. Nonexperimental observational designs, in which investigators do not directly intervene but instead develop methods for describing events that occur naturally and their effect on study subjects, characterize much of effectiveness research. The majority of effectiveness
research studies rely on medical records and related sources such as
claims data collected for billing purposes and hospital discharge
abstract data collected principally for quality assurance purposes.

6. Risk adjustment of patient outcomes is necessary in effectiveness
research to account for the differing risks patients bring to the clinical
setting (e.g., based on patient demographics or conditions un-
related to their presenting illness). Most effectiveness research is
conducted under nonexperimental conditions, and these varying
risks and their potential for confounding evaluations of the per-
formance of a given intervention must be adjusted for in the analysis.

OVERVIEW

The fundamental questions posed in this chapter are (1) What is effec-
tiveness? and (2) How should the effectiveness of healthcare be assessed?
Chapter 3 considers (3) To what extent has effectiveness been achieved?
and (4) What policy strategies contribute to enhancing effectiveness?

A widely accepted definition of effectiveness is the degree to which
improvements in health now attainable are, in fact, attained (Donabedian
2003, 6). Questions of effectiveness have assumed great importance in
recent years because of the continually escalating costs of medical care;
evidence of wide and unexplained variations in the rates of utilization of
medical care across states and regions; community-level evidence sug-
gestig the limited effectiveness of medical care as contrasted to non-
medical factors in improving the health of populations; clinical evidence
of the potential for improvement in the provision of medical care; and
healthcare reform, both at the federal and state levels in the United States.

While the health of the population overall in the United States has
improved substantially over the past century, the health of certain vul-
nerable groups has declined. Specifically, trends make evident the signif-
ificant reduction of acute infectious disease mortality, declines in
mortality from major chronic diseases, and resulting increases in life
expectancy. But, as pointed out in Chapter 1, health status continues
to differ by race as well as by other demographic variables; the differ-
ences are substantial (see Appendix 1). Rates of death due to breast
cancer and cervical cancer are much higher for blacks compared to other races.
In 2001, the death rate per 100,000 was 35.0 in blacks compared to 26.0
and 16.2 in whites and Hispanics, respectively, for breast cancer and
4.9 in blacks compared to 2.3 and 3.4 in whites and Hispanics, respec-
tively, for cervical cancer. In addition, substantial geographic variations
exist in the levels of medical care resources as well as in the rates for
various medical and surgical procedures. These findings raise the ques-
tions of whether health improvements are in fact attributable to med-
care or to some other factor or set of factors, whether the continuing
disparities for selected groups are a result of failures in medical care,
and whether geographic variations are associated with varying outcomes
for patients across areas. These and related questions are addressed by
effectiveness research.

In 1996, the Institute of Medicine (IOM) launched a major initiative
focused on assessing and improving the nation’s quality of healthcare,
known as the Health Care Quality Initiative (IOM 2003). The first phase
of the initiative built on an intensive review of the literature conducted
by the RAND Corporation and a framework that defined the nature of
the quality problem as one of overuse, misuse, and underuse of health-
care services. The report emanating from phase one of the initiative,
Ensuring Quality Cancer Care, documented the wide gulf that exists
between ideal cancer care and the reality that many Americans with
cancer experience (IOM 1999a). During the second phase, spanning 1999
through 2001, the report To Err Is Human: Building a Safer Health System
documented how tens of thousands of Americans died each year from
medical errors (IOM 1999b). A second report from phase two, Crossing
the Quality Chasm: A New Health System for the 21st Century, defined
six aims—care should be safe, effective, patient centered, timely, effi-
cient, and equitable—and related rules for care-delivery redesign to
reduce the magnitude of errors (IOM 2001). Phase three of IOM’s Health
Care Quality Initiative focused on operationalizing the vision of a future
health system described in the Quality Chasm report, emphasizing
reform at three different overlapping levels of the system: the environ-
mental level, the level of the healthcare organization, and the interface
between clinicians and patients.

This chapter presents and discusses a conceptual framework for
effectiveness research as a foundation for evaluating the success of the
U.S. healthcare system in improving the quality of healthcare and ulti-
mately the health of patients and populations. Key methods of effec-
tiveness research from both the clinical and population perspectives are
presented and illustrated using breast cancer screening effectiveness
studies as examples. Chapter 3 categorizes the various policy strategies
for enhancing population health, reviews the evidence on the effectiveness of each strategy, and develops a set of criteria for assessing policy alternatives in terms of effectiveness in the context of breast cancer prevention and treatment.

CONCEPTUAL FRAMEWORK AND DEFINITIONS

Two Perspectives
Effectiveness research reflects two seemingly competing, but in fact complementary, definitions of effectiveness (see Figure 2.1). One represents a population perspective, or macro level view, that considers the role of physical, social, and economic environments on the health of the population. This macro definition is represented in the earlier conceptual work of Milio (1983) and the later work of Evans, Barer, and Marmore (1994). It can be characterized as the epidemiology of health. It includes in its purview both patients who have received medical care and individuals in the population as a whole who have not.

The second is a clinical perspective that represents the micro-level view and focuses on the interactions of patients and providers in the medical care system and institutions and the resulting clinical improvement or health benefits achieved by patients. Research conducted from this point of view examines the impact of the structures and processes associated with delivering medical care in achieving improvements in the health of patients. It is represented in the work of Donabedian (1993); Kerr White and his colleagues, who introduced the concept of the “ecology of medical care” (White 1997; White, Williams, and Greenberg 1961); Wennberg (1990), who has labeled this area “clinical evaluation science”; and Brook and Lohr (1985), who called for an epidemiology of medical care.

Health services research related to each of these views mirrors their differing perspectives.

The epidemiology of health focuses on the benefits from both medical and nonmedical determinants of the health of the population, including environment, behavior, human biology, and medical care. The epidemiology of medical care, or clinical evaluation science, delineates the benefits from medical care for patients. The following example illustrates these differing views.

The National Health and Nutrition Examination Survey, conducted by the National Center for Health Statistics, periodically collects both interview and physical examination data on a representative sample of the U.S. population. Based on the 1999–2000 survey, Hajjar and Kotch (2003) determined that of 1,000 people in the population, 287 had elevated blood pressure (hypertension). Of these 287 people, 168 were under medical care for their hypertension, and 89 had their blood pressure effectively controlled by this treatment. Translating this into the perspectives discussed, from a clinical perspective, 89 of 168, or 53 percent, of hypertensive patients had their blood pressure effectively controlled, but from a population perspective, 89 of 287, or 31 percent, of the hypertensive individuals in the population had their hypertension effectively controlled. The difference between the two views leads to widely varying empirical estimates of primary care effectiveness: 53 percent versus 31 percent.

The differing perspectives also explore different factors to account for the respective rates: biological or clinical factors, or patient adherence, versus social or behavioral factors, or medical care access. A point worth noting, however, is that while this description of two perspectives is instructive, a middle ground may exist between the two perspectives involving health promotion and disease prevention services that encompasses both medical and nonmedical interventions. This middle ground might be labeled health- (in contrast to medical) care and represents the transition to a broader set of policy alternatives for improving the health of the population (see Figure 1.1).
Conceptual Frameworks

The major conceptual frameworks that guide effectiveness research are derived in the population perspective from Evans, Barer, and Marmor (1994) and in the clinical perspective from the work of White, Williams, and Greenberg (1961); Kane (1997); and Donabedian (2003). The Evans, Barer, and Marmor framework, represented in the population perspective displayed in Figure 2.1, defines the determinants of health as the physical and social environment, human biology, individual behavior, and medical care services. Kerr White and his colleagues are responsible for describing what they termed the ecology of medical care, represented within the clinical perspective displayed in Figure 2.1 at the system, institution, and patient levels, but also including the community level as represented by the medical care component of the population perspective displayed in this figure (White 1997; White, Williams, and Greenberg 1961). From the clinical perspective, Donabedian (2003) first offered the categorization of medical care in terms of structure, process, and outcomes for the purpose of determining those aspects that might be indicators of quality. This is shown in Figure 2.1 as the components that can be examined at the patient, institution, and system levels.

Two conceptual models are presented that serve to further clarify and delineate the central determinants of health from the population and clinical perspective, respectively: Kindig and Stoddart's (2003) concept of the field of population health, and Donabedian's (2003) and Kane's (1997) conceptual models for quality and outcomes research. Kindig and Stoddart (2003, 380) have defined population health as “the health outcomes of a group of individuals, including the distribution of such outcomes within the group.” Figure 2.2 displays a schematic definition of the field of population health from the work by Kindig and Stoddart. This framework centrally incorporates both nonmedical and medical determinants of health over the life course and importantly points out the role of policies and interventions at both the individual and social levels to influence these determinants and, ultimately, the level and distribution of health in the population.

A conceptual framework to analyze the effectiveness of clinical prevention and treatment (e.g., for breast cancer) can be divided into the three classic compartments of structure, process, and outcomes (Donabedian 2003; Kane 1997). This framework for assessing effectiveness in the context of the clinical perspective is illustrated in Figure 2.2:

2.3. In this case, structure refers to elements of medical care that are associated with the receipt of services. These may include factors such as availability of insurance coverage and access to facilities. Process refers to the intervention being tested for effectiveness, such as mammography screening. Finally, structure and process together result in outcomes. In the breast cancer example, the intermediate outcome is the early detection of breast cancer, and the final outcome is the improved life expectancy as a result of early detection.

All components of medical care need to be constantly evaluated in the following ways for effectiveness:

1. **Measuring**, for example, the intensity of the intervention (in terms of rates of women obtaining screening), improvement in outcomes (earlier diagnosis of breast cancer), or changes in the structure (increased insurance coverage for screening);
2. **Monitoring** the types of populations receiving the screening, the adoption rates of screening, and improvements in outcomes;
3. **Benchmarking** these aspects of effectiveness against extant standards of quality (such as Healthy People 2010 goals and objectives or...
reside. System refers to the healthcare system, including resources such as money, people, physical infrastructure, and technology and "the organizations and systems or networks of organizations that transform these resources into health services and distribute them to consumers," either within a specific region or for the country as a whole (Longest 2002, 54). It includes all of the elements within the system nationally or in a specific region. Institution refers to a specific organizational entity such as a hospital, clinic, or managed care organization. Patient refers to the recipient of services at the clinical level where the focus is on prevention, treatment, or follow-up and includes an encounter between a patient and a provider. As illustrated in Figure 2.1, each of the three levels within the clinical perspective, as well as the medical care system as a component of the population perspective, can be further elaborated in terms of structure, process, and outcome.

Structure refers to

... the conditions under which care is provided. These include: (1) material resources, such as facilities and equipment; (2) human resources, such as the number, variety, and qualifications of professional and support personnel; (3) organizational characteristics, such as the organization of the medical and nursing staffs, the presence of teaching and research functions, kinds of supervision and performance review, methods of paying for care, and so on. (Donabedian 2003, 46)

Delivery system characteristics as defined in the seminal Medical Outcomes Study (Tarlov et al. 1989) include the organization, specialty mix, workload, and access and convenience of care; provider characteristics of specialty training, preferences, and job satisfaction; and patient characteristics of diagnosis and condition, severity, comorbidity, and health habits. The population perspective would view this component relative to the denominator of a community's population as a whole, while the clinical perspective would lodge it in relationship to the enrollees or patients to be served by a given system or institution.

Process refers to "...activities that constitute health care—including diagnosis, treatment, rehabilitation, prevention, and patient education—usually carried out by professional personnel, but also including other contributions to care, particularly by patients and their families"
(Donabedian 2003, 46). Examples of process variables within the clinical perspective include technical aspects such as visits, medications, referrals, test ordering, and hospitalizations and interpersonal characteristics such as interpersonal manner, counseling, and communication level on the part of patients or enrollees. A population perspective would consider utilization rates or de facto (realized) access for the target population in an area.

Outcomes are "...taken to mean changes (desirable or undesirable) in individuals and populations that can be attributed to health care. Outcomes include: (1) changes in health status; (2) changes in knowledge acquired by patients and family members that may influence future care; (3) changes in the behavior of patients or family members that may influence future health; (4) satisfaction of patients and their family members with the care received and its outcomes" (Donabedian 2003, 46-47).

Examples of clinical outcomes include endpoints such as symptoms and signs of problems, laboratory values, disability, and death and health-related quality of life including the physical, mental, social, and role dimensions. The population perspective would focus on overall population mortality, morbidity rates, or health status.

Structure, process, and outcome are linked conceptually in a research paradigm that assumes structural elements of healthcare as having an influence on what is and is not done in the process as well as how well it is done; this process in turn influences the health outcome people experience as a result of their encounters with the process. This categorization and the implied linkage among these components has become the basic conceptualization in studying the effectiveness of medical care and its determinants. These relationships are captured in the summary framework shown in Table 2.1.

Definitions
Table 2.1 provides several illustrative definitions for the major components of the structure, process, and outcomes of care; the discussion that follows summarizes the key idea of each. Quantity refers to the number of physicians, nurses, and other providers as well as the quantity of monetary resources. Efficacy is concerned with the benefits achievable from a therapy or an intervention under ideal conditions, such as those found in a randomized clinical trial (Cochrane 1971; Donabedian 2003; Lohr, Eleazer, and Mauskopf 1998; Sackett 1980; Williamson 1978).

Variations in use relate to the quantity, or what is more commonly referred to as utilization, of healthcare services and procedures (Lohr, Eleazer, and Mauskopf 1998). It also includes the frequency or volume of procedures done. Quality is an attribute of the healthcare process having to do both with whether the right thing is done and whether it is done well (Brook and Lohr 1985; Donabedian 1973, 1980, 1982, 2003; Lohr, Eleazer, and Mauskopf 1998). The Institute of Medicine (2003) Health Care Quality Initiative has defined quality as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge." Quality assessment thus deals with evaluating the process of healthcare in the service of ultimately improving health outcomes. Appropriateness is the subset of quality that concerns determining whether the right thing was done for the patient.

Effectiveness concerns the results achieved in the actual practice of healthcare with typical patients and providers, in contrast to efficacy, which is assessed by the benefits achieved under ideal conditions (Cochrane 1971; Donabedian 2003; Lohr, Eleazer, and Mauskopf 1998; Sackett 1980; Williamson 1978). Quality is that part of the gap between efficacy, or what is achievable, and effectiveness, or what is achieved, that can reasonably be attributed to healthcare itself. Evidence-based medical care focuses on the use of the best available efficacy and effectiveness evidence to inform decisions about patient care and guide healthcare policy (Lohr, Eleazer, and Mauskopf 1998).

In conclusion, the argument has been made that both the population and clinical perspectives are important in examining the effectiveness of healthcare. The population perspective argues that nonmedical as well as medical investments are required to improve the health of individuals and communities, while the clinical perspective illuminates how enhancing the precision of medical care can contribute to this improvement.

KEY METHODS OF ASSESSING EFFECTIVENESS
Two basic questions related to effectiveness were presented at the beginning of this chapter: (1) What is effectiveness? and (2) How should the effectiveness of healthcare be assessed? Table 2.2 introduces a framework for effectiveness research that attempts to integrate the two perspectives—population and clinical—and the four levels—community, system, institution, and patient—in empirically addressing these questions. The
<table>
<thead>
<tr>
<th>Structure</th>
<th>Process</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>Variations in Use</td>
<td>Effectiveness</td>
</tr>
<tr>
<td>Efficacy</td>
<td>—Quantity</td>
<td>—Mortality</td>
</tr>
<tr>
<td></td>
<td>—Quality</td>
<td>—Morbidity</td>
</tr>
<tr>
<td></td>
<td>—Appropriateness</td>
<td>—Health status</td>
</tr>
</tbody>
</table>

**QUANTITY**
Refers to the number of physicians, nurses, and other providers as well as the quantity of monetary resources.

**VARIATIONS IN USE**
Refers to different observed levels of per capita consumption of a service, especially hospital care, office visits, drugs, and specific procedures.

**EFFECTIVENESS**
Refers to actual achieved benefit.

**Does it work?**
Does the maneuver, procedure, or service do more good than harm to those people to whom it is offered?

---

**EFFICACY**
Refers to maximum achievable benefit.

**Can it work?**
Does the health maneuver, procedure, or service do more good than harm to people who fully comply with the associated recommendations or treatment?

This question refers to the ability of a particular medical action in altering the natural history of a particular disease for the better, under ideal conditions.

**QUALITY**
Refers to a judgment concerning the process of care, based on the extent to which that care contributes to valued outcomes.

Quality of medical care is that component of the difference between efficacy and effectiveness that can be attributed to care providers, taking account of the environment in which they work.

**APPROPRIATENESS**
Refers to the extent to which available knowledge and techniques are used or misused in the management of illness and health.

This question refers to the ability of a particular medical action in altering the natural history of a particular disease for the better, under actual conditions of practice and use.

**Sources:**
- Cochrane (1971);
- Donabedian (2003);
- Lohr, Eleazer, and Mauskopf (1998);
- Sackett (1980);
population perspective focuses on addressing these questions in the context of a community-level analysis, while the clinical perspective can seek to address them at the system, institution, or patient level of analysis or at a combination of these levels.

The outcome measures, risk-adjustment procedures, study designs, and data sources that might be drawn on at each level, as well as examples that illustrate the application of these methods at each of the levels, are highlighted in Table 2.2 and are discussed in the sections that follow.

**Outcome Measures**

Outcome measures fit into the general categories of mortality, morbidity, and health status, although specific indicators of each of these broad outcome measures may be used at different levels. The usefulness of a measure depends in part on the degree to which it meets the criteria of reliability, validity, feasibility, and sensitivity (McDowell and Newell 1996). **Reliability** concerns the reproducibility of the measure under various conditions of administration. **Validity** relates to the accuracy of the measure, in the sense that it measures what it is intended to measure. **Feasibility** refers to the ease with which the scale can be used in various populations. **Sensitivity** refers to the ability of the measure to detect changes—improvement or deterioration—in the condition of the person as a result of healthcare. Sensitivity to healthcare variation is particularly important for outcome measures being used to assess the effectiveness of care.

Community-level outcome measures include population death, morbidity, and disability rates as well as disease prevalence and incidence rates and perceived health status. One problem with these measures is how to combine them into a positive index of the community’s or population’s health, as opposed to negative indexes such as death rates, to yield a representation of health-adjusted life expectancy (Kindig 1997). The **disability-adjusted life year (DALY)** represents one attempt to combine these community-level measures in a way that reflects the burden of disease on a population (Murray and Lopez 1996). Specifically, DALY expresses years of life that are lost to premature death and years lived with a disability of specified severity and duration. One DALY is thus equivalent to one lost year of healthy life. In a comprehensive review, the proponents of this approach have examined the global burden of disease using the DALY measure. They find that the leading causes of disease burden were the following: childhood and maternal underweight (138 million DALYS, 9.5 percent), unsafe sex (92 million DALYS, 6.3 percent), high blood pressure (64 million DALYS, 4.4 percent), tobacco (59 million DALYS, 4.1 percent), and alcohol (58 million DALYS, 4.0 percent) (Ezzati et al. 2002). Another attempt to index a population’s health was based on the years of healthy life derived from a combination of the responses to the activity limitation and self-perceived health status questions from the National Health Interview Survey (NCIHS 1995).

Population-based data, as exemplified by mortality rates, while relatively high in reliability, validity, and feasibility, have been shown repeatedly to be insensitive to medical care variation. This suggests that these data are useful for addressing the question of medical care’s contribution to the health of the population, but they have limited applicability in assessing the clinical effectiveness of medical care.

For examining mortality, morbidity, and health-status outcomes across institution and system levels, outcomes may be aggregated from the patient level within these categories. Patient-level outcome measures focus on individual deaths, on morbidity as reflected both in adverse outcomes and disability limitations, and on the health status outcomes of clinical endpoints, such as blood sugar and blood pressure and subjective health-status measures. At the institution and system levels, these patient-level measures are aggregated to produce case fatality, complication, disability, and diagnosis rates as well as averaged subjective health status for groups of patients.

**Subjective Health-Status Measures**

Subjective health-status (SHS) measures, based on individuals’ self-reports, are singled out here for a more detailed discussion because they have undergone extensive development and have achieved widespread use in the past two to three decades (Ware 2003). They may be of a generic type, applicable across all disease conditions, or of a disease-specific type. Both types are needed in the assessment of the medical care outcomes—the generic indicators for comparisons across disease conditions, and the disease-specific ones to more sensitively identify the effects of diseases on people and the effects of treatments on particular disease conditions. A range of possible SHS measures has been collected in Measuring Health: A Guide to Rating Scales and Questionnaires (McDowell and Newell 1996). For each of over 100 instruments, this book presents a description, copies of the actual questionnaire, information on the reliability and validity of the instrument, and a
<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Population Perspective</th>
<th>Clinical Perspective</th>
<th>Institution</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Community</td>
<td>System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td></td>
<td>Mortality</td>
<td>Mortality</td>
<td></td>
</tr>
<tr>
<td>Population death rates</td>
<td></td>
<td>Case fatality rates</td>
<td>Case fatality rates</td>
<td></td>
</tr>
<tr>
<td>Morbidity</td>
<td></td>
<td>Morbidity</td>
<td>Morbidity</td>
<td></td>
</tr>
<tr>
<td>Population morbidity rates</td>
<td></td>
<td>Complication rates</td>
<td>Complication rates</td>
<td></td>
</tr>
<tr>
<td>Disability rates</td>
<td></td>
<td>Disability rates</td>
<td>Disability rates</td>
<td></td>
</tr>
<tr>
<td>Health status</td>
<td></td>
<td>Health status</td>
<td>Health status</td>
<td></td>
</tr>
<tr>
<td>Disease incidence and prevalence rates</td>
<td></td>
<td>Diagnosis rates</td>
<td>Diagnosis rates</td>
<td></td>
</tr>
<tr>
<td>Perceived health status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study designs</td>
<td>Observational-epidemiological</td>
<td></td>
<td>Averaged HRQOL</td>
<td>Health-related QOL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Demographic characteristics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comorbidity rates</td>
<td>Comorbidity rates</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk-adjustment systems</td>
<td>Risk-adjustment systems</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Observational-</td>
<td>Observational-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>interorganizational</td>
<td>interorganizational</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk adjustment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study designs</td>
<td>Observational-epidemiological</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data sources

- Records
  - Population health information system
  - Vital statistics
  - Disease surveillance
- Surveys
  - Chinese American community screening (Tu et al. 2003)
  - European national screening program (de Koning 2000)

Example

- Records
  - Medical records
  - Discharge data
  - Claims data
- Surveys
  - Public hospital clinic screening program (Thompson et al. 2000)
- Surveys
  - Patient screening in response to intervention (Ell et al. 2002)

Typical Effectiveness Research Questions by Level of Analysis

**Community**

- What is the contribution of medical care to the health of the population?

**System**

- What is the impact of system-level variables (e.g., provider specialty mix, organizational form, payment mechanism) on the processes and outcomes of medical care?

**Institution**

- What is the impact of the quality of care on the outcomes of medical care?

*Note: HRQOL = health related quality of life; QOL = quality of life; RCT = randomized clinical trial.*
complete listing of references. McHorney (1999) also provides a brief summary and review of the ten most widely used SHS measures.

One example of generic SHS measures is the Medical Outcomes Study Short Form (SF-36), with its simpler companion, the SF-12. This questionnaire contains an array of questions related to the effect of disease on physical health, mental health, and social health, as well as on health perceptions. The scores on each of eight dimensions were originally kept separate and presented as a profile. They have subsequently been grouped into two categories to yield summary scores for physical and mental health, respectively (QualityMetric 2003; Ware, Kosinski, and Keller 1996).

In addition to the generic measures are disease-specific subjective health status measures, such as the European Organization for Research and Treatment of Cancer Core (Aaronson et al. 1993) and the Functional Assessment of Cancer Therapy—General (Cella et al. 1993), which have been developed specifically for cancer patients, and the Dermatology Life Quality Index (Finlay and Khan 1994) and the Skindex (Chren et al. 1997), which have been developed for patients with skin disease. Such questionnaires focus on unique aspects of the disease for which they were developed. All the instruments described above are reliable; valid; feasible; and, above all, sensitive to the changes following medical treatment.

The optimum strategy for outcomes assessment, given the different levels of sensitivity, may be to use a generic instrument, such as the SF-36 or the SF-12, supplemented with disease-specific questions or a disease-specific questionnaire (e.g., see Ren et al. 1998). The generic instrument would allow comparisons to be made across diseases, while the disease-specific instrument would be more likely to provide sufficient sensitivity to detect small changes in patients' conditions.

**Other Subjective Outcome Measures**

In addition to the above measures, the use of other patient-centered outcome measures to evaluate aspects of medical care quality such as satisfaction (which evaluates perceptions of discrete past healthcare transactions) and trust (a forward-looking assessment of an overall relationship) with medical care is becoming commonplace in many health services research studies. Several measures of trust of and satisfaction with physicians, insurers, and the medical profession have been developed and validated and are being used in assessments of overall healthcare quality (Balkrishnan et al. 2003). Consumer assessments of healthcare provide important information about how well health plans and clinicians meet the needs of the people they serve. The Agency for Healthcare Research and Quality has sponsored development of the Consumer Assessment of Health Plans Survey (CAHPS), an integrated set of tested and standardized questionnaires and reporting formats to collect and report information about the experiences of consumers enrolled in health plans since 1995 (Hays et al. 1999). Studies have reported that CAHPS ratings could affect consumer selection of health plans and ultimately contain costs (Spranca et al. 2000).

**Risk Adjustment**

Ultimately, the value of all effectiveness research depends on the ability to validly adjust for differences in risks associated with final outcomes. Risk adjustment of patient outcomes is necessary in effectiveness research to account for the differing risks patients bring to the clinical setting. Clearly, patients who differ at admission to a hospital in their risks and who receive similarly effective treatments will experience different outcomes. When randomized clinical trials are possible, these differences can be minimized by the random allocation of subjects to experimental and control groups. But under nonexperimental conditions, under which most effectiveness research is done, these differences and their potential confounding should be adjusted for in the analysis. These differing risks that require adjustments include differences in patient demographic characteristics such as age, gender, and race; comorbidities, or conditions unrelated to the primary illness that can negatively affect treatment outcomes; and diagnoses that may differ in the initial severity of illness.

At the patient level, two general approaches may be taken in this adjustment. A subjective approach, relying on the informed judgment of experienced clinicians in rating the severity of the patient's illness at entry, may provide a valid assessment of a patient's status (Charlson et al. 1986), but such an expensive procedure, in terms of the physician's time, is rarely possible. In its place, an objective approach, utilizing clearly identified data related to the patient's risk, clinical state, and probable outcome, applies an algorithm or formula to generate a score characterizing the patient's risk. These data may include characteristics
of patients, their comorbid conditions, and their diagnoses, which at
the institution and system levels may be incorporated into risk-adjust-
ment systems such as the Acute Physiological and Chronic Health
Evaluation (APACHE) scale or the Medical Illness Severity Grouping
System (MedisGroups) described later. Also at the institution and sys-
tem levels, where some of the detailed patient data such as discharge
and claims data may be lacking, demographic characteristics or com-
orbidity rates may be used as proxies for actual severity measures. At
the community level, demographic characteristics such as age and gender
are used to adjust for differing risk of illness.

The risk-adjustment methods in common use assume the objective
approach described above; n types of them, all at the patient level, are
described and thoroughly analyzed in Risk Adjustment for Measuring
Health Care Outcomes (Iezzoni 2003). This resource also provides in-
formation on the dimensions of risk, data sources, and performance of
these measures. Iezzoni specifies several issues important to the as-
sessment and measurement of risk: level of analysis, time frame for ob-
server, timing of data collection, feasibility, reliability, and validity. The
level or unit of analysis determines both the data that are available and
the dimensions important to consider. Health-related quality of life
(HRQOL), for example, is not likely to be found when entire systems are
being analyzed, because HRQOL is not routinely collected on all patients
in all settings. The time frame for observation of outcomes—for ex-
ample, whether consideration is to be given only to hospital inpatient events
or to things occurring within six months postdischarge—determines
which dimensions are important. The timing of data collection is also
important; if the severity of illness at admission is to be the basis for the
risk adjustment, it is not appropriate to use risk data gathered from the
entire stay in the hospital because the results may be confounded by the
treatments subsequent to admission.

The n different systems for risk adjustment that Iezzoni (2003)
describes represent a mix of dimensions as well as disease-specific and
generic measures. All are proprietary to some extent, and are therefore
less available for the kind of critical analyses done of outcome mea-
sures. Two of these, APACHE and MedisGroups, will be described now
in greater detail. Both are grounded in the clinical perspective on effec-
tiveness. APACHE, one of the first risk-adjustment systems developed,
continues to be updated and widely applied. MedisGroups has been
mandated in several states as the system to be used by hospitals. (See
Cardinal Health [2003] and Cerner [2003] for more information on
the vendors for these systems.)

The APACHE scale was developed for the evaluation of patients in
hospital critical care units and uses a dozen physiologic values gener-
ated from physical findings and laboratory data in the first 24 hours
after admission (Knaus et al. 1985). Scale scores for each of these val-
ues are added and are combined with adjustments for age and chronic
health conditions to yield an overall score. The APACHE score is generic
in that it can be applied across diseases for comparisons of severity.
Subsequent versions of APACHE, such as APACHE III, added several more
variables to the risk-adjustment scale such as age and additional comor-
bid conditions (Knaus et al. 1991; Knaus 2002).

MedisGroups has assumed great importance because several states
have mandated it as the risk-adjustment measure to be used by hospi-
tals in reporting data to state agencies. The MedisGroups system pro-
duces a generic set of severity categories across illnesses, using medical
record data processed by a proprietary program (Brewster et al. 1985).
It uses “key clinical findings,” including laboratory, radiology, pathology,
and physical examination data. This information can be input into the
system’s coding scheme to permit the severity of patients’ condi-
tions at admission, as well as their progress over the course of the hos-
pitalization, to be monitored. The MedisGroups system has been widely
used for risk adjustment in hospital-based health services research stud-
ies (Iezzoni et al. 1998; Silber et al. 1999).

These risk-adjustment methods and severity-of-illness measures can
and should be evaluated by the same criteria as health-status measures
—reliability, validity, feasibility, and sensitivity. Reviews (Iezzoni 2003)
and studies (Hwang et al. 2001) have summarized the information from
studies comparing measures of these attributes. Because the objective
methods use factual data and are computerized, they are both basically
reliable and feasible. Their validity continues to be a question, and they
vary in sensitivity. Risk adjustment is less well developed than outcome
measurement and is impeded by the fact that much of the work is in
the proprietary domain. Iezzoni, however, has contributed greatly to
removing the veil from these systems.
Study Designs

Study designs for effectiveness research cover a range of possibilities (see Table 2.2). The design principles are the same as those for any study: maximize experimental variance, control extraneous variance, and minimize error variance. On one hand, applying these principles results in outcomes research designs that follow true experimental design principles of random allocation, control groups, blinding, and homogeneity and lead to efficacy studies. On the other hand are nonexperimental observational designs, in which investigators do not directly intervene but instead develop methods for describing events that occur naturally and their effect on study subjects. These types of studies characterize much of effectiveness research and are represented by examples of effectiveness studies presented later in this chapter. Alternatives include meta-analysis and decision analysis, sometimes called synthetic designs (Petitti 1999).

The assessment of efficacy—the determination of the benefits of a particular medical therapy, health service activity, or public health intervention under ideal conditions—involves the randomized clinical trial (RCT) as the primary method of analysis. The hallmark of the RCT is random assignment of patients to experimental and control groups and, hence, control of much of the extraneous variation and sources of error. Good examples of RCTs concerning medical therapies include trials of major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker versus diuretics; the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT 2002) and the Women's Health Initiative Memory Study, a trial of the effect of estrogen therapy in preventing and slowing the progression of dementia (Shumaker et al. 2003). Large, simple trials have been used to extend the range of RCTs to situations where treatment effects are modest in size (Peto, Collins, and Gray 1995).

RCT designs have been used to assess the benefits of ways in which medical care is delivered, for instance, in the evaluation of effectiveness of pharmacist care for patients with reactive airways disease (Weinberger et al. 2002) and in the comparison of sentinel-node biopsy with routine axillary dissection in breast cancer (Veronesi et al. 2003). This design was also used to assess the effects of different medical care payment plans on use and outcomes in the RAND Health Insurance Experiment (Brook et al. 1983). RCTs have been used less frequently to evaluate public health interventions (Orleans 1995); the polio vaccine trials conducted in the 1950s are, however, an outstanding exception and a good example of large public health trials.

RCTs are a powerful method but are not often feasible because of the expense necessitated by the large sample sizes and the length of time required to conduct them. Ethical issues involved in depriving patients of treatment may also preclude their use. Another problem with RCTs is that many are done on a small, local scale, and therefore the results often are not highly significant nor widely disseminated. Because of the difficulties and expense inherent in RCTs, they must be used selectively. As a result, many important treatment questions cannot be answered by such designs.

An alternative to RCTs is the use of synthetic methods such as meta-analysis and decision analysis. Meta-analysis involves quantitatively and statistically combining the results of several RCTs to estimate the results of therapy when no single trial may be sufficient in number of patients to yield a statistically significant result (Petitti 1999). Meta-analyses have been used, for example, to obtain estimates of the effectiveness of nocturnal, noninvasive, positive-pressure ventilation in patients with stable chronic obstructive pulmonary disease (Wijkstra et al. 2003), efficacy and safety of ephedra and ephedrine for weight loss and athletic performance (Shekelle et al. 2003), and health outcomes associated with antihypertensive therapies used as first-line agents (Psaty et al. 2003).

Meta-analyses are also being used in the Cochrane Collaboration (2003), which is a major international effort directed to ensuring that all areas of healthcare that have been evaluated using RCTs would be covered. Collaborators in this process prepare and maintain systematic reviews of RCTs as well as other evidence where appropriate. These reviews are then maintained in a database and disseminated for use.

Decision analyses synthesize information about effectiveness to determine the value of one approach versus another for policy analysis and ultimately for clinical decisions. A decision analysis requires information on the actual treatment of patients with disease, the outcomes, and the value of those outcomes to patients. Information from large databases and other sources is used to estimate the probabilities of different outcomes from therapy for patients. Patient surveys provide information on patients' symptoms as well as their preferences for different outcomes. The advantage of a decision analysis is that it synthesizes a large amount of information relevant to effectiveness. The disadvantage
is that necessary data on patient values or preferences are often not available. An example of decision analysis using national-level Medicare data concerning breast cancer screening (Mandelblatt et al. 1992) is discussed in Chapter 9.

Observational nonexperimental designs, chiefly cross-sectional studies, are another design type, and characterize the bulk of the literature on effectiveness. Community-level epidemiological studies use existing registries or databases such as the Medicare database. System- and institution-level studies use databases appropriate to these levels, such as hospital discharge and claims data, to address issues within the organization (intraorganizational) and between organizations (interorganizational). At the patient level, when rcts are not possible, reports of individual cases and series of them are reported. One problem with these observational studies is that the databases used may have significant biases due to missing values, lack of validating evidence, or the unrepresentativeness of the database and that these biases may limit the generalizability of study results. Another problem is that random allocation has not been possible, and therefore statistical adjustments are applied to account for the differences in mix of patients that may result from possible selection biases. Despite these limitations, however, observational studies are more feasible than experimental studies and draw on large volumes of data reflecting actual experience. Most of the studies of effectiveness, and thus much of the evidence about effectiveness to be presented and reviewed in the next section of this chapter, derive from these observational designs.

Data Sources

Where clinical trials and true experiments are not possible, effectiveness research relies on a variety of data sources for the cross-sectional studies (as shown in Table 2.2). Surveys of institutions, providers, and patients, as well as records of medical care, discharges, and claims, provide information for effectiveness research at the patient, institution, and system levels, dependent on the level of aggregation of the data. At the community level, public health surveillance systems and vital statistics data that may be used to construct population health information systems provide the data for effectiveness studies.

Community-level outcome measures such as population mortality rates can be obtained from state and U.S. Vital Statistics data, as well as from the World Health Organization and the Organization for Economic Cooperation and Development data for international comparisons. Morbidity rates can be acquired from the National Center for Health Statistics (nhis) National Health Interview Surveys (NHIS). These surveys are conducted annually on a sample of the U.S. population and yield, among other data, information on limited activity days and restricted activity days. Disease prevalence and incidence data can also be obtained from NHIS data, as well as from other federal sources such as the Mortality and Morbidity Weekly Report of the Centers for Disease Control and Prevention.

The surveys conducted by NHIS provide a rich source of information for effectiveness research. Information is gathered, for example, by the NHIS on respondents’ use of medical care. Provider surveys, such as the National Ambulatory Medical Care Survey, the National Hospital Discharge Survey, and the National Long-Term Care Survey, provide aggregate information on patients and their use of healthcare services. Diseases are also recorded in national registries, some maintained by the government and some by private sources. An example of the former is the National Cancer Institute’s Surveillance Epidemiology and End Results Program (2003) database.

The majority of effectiveness research studies, however, rely on medical records and related sources such as claims data collected for billing purposes and hospital discharge abstract data collected principally for quality assurance purposes. Administrative databases are maintained to keep records of delivery of healthcare services, such as reimbursing medical care providers or determining patient eligibility for certain services covered in part or full by an insurance provider. This information is collected and maintained by the payer (government or private insurer) primarily for reimbursement, but also for other primarily financial reasons such as monitoring, enrollee risk assessment, and rate setting.

The three major sources of these databases are the federal government (Medicare, VA), state governments (Medicaid), and private insurance companies (HMOs, PPOs). Other large databases such as the Health Plan Employer Data and Information Set (HEDIS) and the Computerized Needs-Oriented Quality Measurement Evaluation System (CONQUEST) have been developed for evaluation of quality and performance of healthcare organizations. The type of data collected in each of these databases varies, but most of the administrative databases today contain information on select patient demographics (maintained in an eligibility or enrollment file) and major healthcare service utilization (hospitaliza-
tions, emergency department visits, outpatient physician visits, surgical procedures). Most of these are recorded using several coding systems. The International Classification of Diseases, 9th revision (ICD-9) disease codes are most commonly used to identify the diagnosis, while a number of methods are used to classify medical procedures.

Hospitals usually use the ICD-9 procedure codes, while physicians employ the Centers for Medicare & Medicaid Services Healthcare Common Procedure Coding System, which is largely based on the American Medical Association's Current Procedural Terminology. Some payers also include prescription drug benefit as a part of the insurance, and information on this is recorded in separate outpatient pharmacy files. Medications are usually coded using the National Drug Codes (NDCs). Most of the data in these claims files exists as payment records for each service rendered. The most important structural attribute that determines the utility of the claims data for research in patient outcomes is the ability to link these service records at the individual patient level and construct “episodes of care” per patient. This is usually achieved through a common patient identification number, which is present in all the data files.

However, presence of this patient identifier (usually social security number) also raises serious concerns about patient confidentiality and privacy issues, and these numbers are usually scrambled or replaced before administrative data are released for research. These requirements have become even more stringent in light of the Health Insurance Portability and Accountability Act enforcements. However, the linkage back to the patient may be required for many effectiveness studies, and in some specific types of research where patient identification becomes necessary, full review from the Human Subjects Committee or an institutional review board may be required. Once episodes of care are created for patients, it is possible to track down serious medical events, adverse events requiring secondary or tertiary medical care, and patient mortality types of outcomes data that may be required in effectiveness studies.

The advantages provided by administrative databases include providing information over long periods of time (longitudinal records), as well as more detail regarding procedures and services received than respondents would be able to recall in surveys. Problems with use of administrative data include patients dropping out because they are no longer eligible for insurance; missing data for variables of interest; lack of capture of variables of interest; selection of specific types of patients into insurance plans or provider groups, thus introducing bias; and logistical limitations due to extremely large sample sizes of patients (Lezoni 2003; Quam et al. 1993). Also, one has to remember that the information provided by these databases cannot be used to establish causal temporality; rather, these observational data only imply associations between variables.

Examples
Several studies are presented as prototypical examples illustrating each level of effectiveness research—community, system, institution, and patient (see Table 2.2). They also demonstrate the use of the various effectiveness research methods discussed earlier (i.e., outcomes measured, risk adjustment used, basic study design, and data sources). The prototypical examples that will be reviewed include a community level example of the effectiveness of a breast cancer screening program among women in the Chinese-American community in one U.S. city, a system level example of effectiveness of breast cancer screening programs in European countries, an institution-level example of an effectiveness evaluation of an intervention to increase mammography utilization in an inner-city public health hospital, and finally, a patient-level example involving evaluation of a targeted, structured intervention designed to reduce barriers to diagnostic follow-up adherence and initiation of treatment in low-income women with abnormal mammogram screens.

Community-Level Example
A study by Tu et al. (2003) assesses the current screening knowledge and practices related to mammography screening behavior of Chinese American women in Seattle, Washington, in 1999. The study sample was based on a cross-sectional, complex representative sampling design (community-level), and all interviews were conducted in the respondents’ home by bilingual, bicultural, Chinese American female interviewers. Outcome variables included recent/past receipt of screening mammograms. Specific associations between language concordance with physician, physician ethnicity and gender, and outcome variables were examined. Seventy-four percent of the surveyed women reported prior mammography screening, and 6% percent reported receiving a screening in the past two years. Although language concordance with physician was associated with higher screening rates, similar to the general
population, a recommendation by primary care physician or nurses increased rates of mammography significantly, irrespective of language concordance. Based on this community survey study, the authors recommend a multifaceted approach to increase mammography screening by Chinese American women. This includes aggressive recommendation by the primary care provider and targeted education to address the effectiveness of screening mammography compared to other options (breast self-examination and clinical breast exams) in this specific population.

**System-Level Example**

A review article by de Koning (2000) assessed the potential effectiveness of a national breast cancer screening program that had been instituted in the Netherlands in 1987 in the context of similar attempts in other European countries. In particular, the driver of the national screening program in the Netherlands (system-level) was the data from three large community trials in Sweden as well as national programs in England and Wales that have estimated a 20 percent reduction in breast cancer mortality at the population level. The program comprised two screenings per year for all women between 50 and 70 years. As far as the Dutch program was concerned, at the time of the report’s publication, it was still too early too reach any conclusions about mortality reductions. However, the first short-term results of the screening program had exceeded expectations with 1.4 million of the 1.7 million who had been offered the program being examined and early-stage cancer detection rates of almost 6 out of every 1,000 women screened for the first time. The author calls for extra attention that will have to be given to characteristics of tumors that will be detected during subsequent screens in the Dutch program by comparison with findings in other countries to evaluate the overall cost-effectiveness of maintaining this program in future years.

**Institution-Level Example**

To increase rates of mammography among low-income, urban women in their 50s and 60s, Thompson et al. (2002) conducted a comprehensive (institution-level) intervention in a public hospital delivering comprehensive medical services to low-income residents of a large inner city. A total of 196 eligible women age 50–74 years who were enrolled in the internal medicine clinic, were noncompliant with mammography screening, and had at least one routine clinic appointment during the 15-month study were entered into an RCT of a comprehensive nurse-administered motivational intervention to increase mammography rates. Overall, 49 percent of the women who received the intervention had a mammography within eight weeks of an index visit compared to 22 percent of the control women. There was an additional cost of $151 (1996 U.S. dollars) associated with receiving the intervention itself and $559 in additional cost incurred for each woman who was motivated to receive a mammogram because of the intervention. The societal perspective was used in estimating costs, meaning that the costs included those borne by the patient, the payer (insurer), and society combined. Through this study the authors demonstrated the effectiveness of the motivation program at the institutional level and developed a cost-tracking model while intervening in a clinic institution setting, thus allowing the institution to make informed decisions about implementing programs to increase the motivation of their patients to receive screening.

**Patient-Level Example**

A study by Ell et al. (2002) involved evaluation of a targeted, structured intervention designed to reduce the number of known barriers to diagnostic follow-up adherence and initiation of treatment among low-income women with abnormal mammogram findings in Los Angeles and New York. The intervention consisted of identification of potential barriers through a scripted, structured telephone interview. Based on symptoms reported in the interview, the women were assigned to a risk level for nonadherence based on an empirical algorithm. High-risk patients (patients with significant mental health symptoms, patients with psychosocial stressors, and women who had received a diagnosis of cancer) were then referred to the team social worker for further individual (patient-level) assessment and intervention. Patients also received reinforcing telephone calls every six months. The observational pilot study in 605 women in two large, urban diagnostic centers showed that adherence rates through diagnostic resolution and initiation of treatment in the intervention group was more than 90 percent. Rates of adherence among women who could not be located or who refused study consent was significantly lower (70 percent). The study results supported the combining of interventions and the practical utility of
the clinical decision-making algorithm in determining individualized risk of nonadherence, thereby effectively identifying "at-risk" subjects for further intervention.

SUMMARY AND CONCLUSIONS
In this chapter, effectiveness has been defined in terms of two complementary views—a population perspective and a clinical perspective. The population view asks what contributions medical care makes to the health of the population. The clinical view, by contrast, asks how medical care improves the health of patients who enter the system for care. This chapter presents a framework that attempts to integrate these views in identifying the medical and nonmedical determinants of population health. The key methods of effectiveness research that help provide answers to these questions have been described, discussed in terms of their strengths and weaknesses, and illustrated in a set of example studies. The next chapter illustrates the useful application of these methods in a broad range of outcomes research answering the basic effectiveness questions.

REFERENCES


