Evaluation is a part of everyday life. Does Ford make a better truck than Chevrolet? What kind of reviews did a new movie get? What are the top 10 football teams in the country? Who will be the recipient of this year’s outstanding teacher award? All these questions entail judgments of merit, or “worth,” reached by weighing information against some explicit or implicit yardstick (Weiss, 1972, 1998a). When judgments result in decisions, evaluation is being performed at some level (Shortell & Richardson, 1978).

This book is about the evaluation of health programs and the role it plays in program management and decision making. All societies face sundry health problems. Accidents, cancer, diabetes, heart disease, HIV infection, suicide, inequitable health and access to health care across social groups, and many others are mentioned commonly in the health literature (U.S. Department of Health and Human Services, 2015a). A health program is an organized response, or “intervention,” to reduce or eliminate one or more problems by achieving one
or more objectives, with the ultimate goal of improving the health of society or reducing health inequalities across social groups (Shortell & Richardson, 1978). Interventions are defined broadly and include intentional changes in health systems or other societal institutions to improve individual and population health and reduce health inequalities by socioeconomic status, race/ethnicity, religion, age, gender, sexual orientation or gender identity, mental health, disability, geographic location, or other characteristics linked historically to discrimination or social exclusion (U.S. Department of Health and Human Services, 2015a).

**Evaluation** is the systematic assessment of a program's implementation and consequences to produce information about the program's performance in achieving its objectives (Weiss, 1998a). In general, most evaluations are conducted to answer two fundamental questions: Is the program working as intended? Why is this the case? Research methods are applied to answer these questions, to increase the accuracy and objectivity of judgments about the program's success in reaching its objectives, and to search for evidence of unintended and unwanted consequences. The evaluation process fulfills this purpose by defining clear and explicit criteria for success, collecting representative evidence of program performance, and comparing this evidence to the criteria established at the outset. Evaluations help program managers understand the reasons for program performance, which may lead to improvement or refinement of the program. Evaluations also help program funders to make informed judgments about a program's worth, which may result in decisions to extend it to other sites or to cut back or abolish a program so that resources may be allocated elsewhere. In essence, evaluation is a management or decision-making tool for program administrators, planners, policymakers, and other health officials.

From a societal perspective, evaluation also may be viewed as a deliberate means of promoting social change for the betterment of society (Shortell & Richardson, 1978; Weiss, 1972). Just as personal growth and development are fundamental to a person's quality of life, so do organizations and institutions mature by learning more about their own behavior (Shortell & Richardson, 1978). The value of evaluation comes from the insights that its findings can generate, which can speed up the learning process to produce benefits on a societal scale (Cronbach, 1982).

Evaluation, however, can be a double-edged sword. The desire to learn more is often accompanied by the fear of what may be found (Donaldson et al., 2002). Favorable results typically are greeted with a sigh of relief by those who want the program to
succeed. By contrast, unfavorable results may be as welcome as the plague. When an evaluation finds that a program has not achieved its objectives, the program's very worth is often brought into question. Program managers and staff may feel threatened by poor evaluation results because they often are held accountable for them by funders, who may decide the program has little worth. In this case, funders or other decision makers often have the power and authority to change program implementation, replace personnel, or even terminate the program and allocate funds elsewhere.

For popular health programs, such as prenatal care for low-income women, program advocates may view unfavorable results as a threat to the very life of the program. To a great degree, the worth of prenatal care programs is grounded on the argument that public spending now will prevent future costs and medical complications associated with low birth weight (Huntington & Connell, 1994). Previous evaluations reported “good” news: Prenatal care pays for itself (for every $1.00 spent, up to $3.38 will be saved). The “bad” news is that the evaluations have serious methodologic flaws that may have resulted in overestimates of the cost savings from prenatal care (Huntington & Connell, 1994). Today, the evidence is insufficient to conclude that universal prenatal care prevents adverse birth outcomes and is cost saving (Grosse et al., 2006; Krans & Davis, 2012). These findings have attracted national attention because they challenge the very worth of prenatal care programs if the objective of those programs is to save more than they cost (Kolata, 1994).

In all evaluations, a program’s worth depends on both its performance and the desirability of its objectives, which is always a question of values (Kane et al., 1974; Palumbo, 1987; Weiss, 1983). For prenatal care and other prevention programs, the real question may not be, How much does this save? but more simply, How much is this program worth? (Huntington & Connell, 1994). For health and all types of social programs, the answer to this fundamental question can have far-reaching consequences for large numbers of people. Greenhalgh and Russell (2009) capture the essence of the values quandary in evaluation:

Should we spend limited public funds on providing state-of-the-art neonatal intensive-care facilities for very premature infants? Or providing “Sure Start” programs for the children of teenage single mothers? Or funding in vitro
fertilization for lesbian couples? Or introducing a “traffic light” system of food labeling, so even those with low health literacy can spot when a product contains too much fat and not enough fiber? Or ensuring that any limited English speaker is provided with a professional interpreter for health-care encounters? Of course, all these questions require “evidence”—but an answer to the question “What should we do?” will never be plucked cleanly from massed files of scientific evidence. Whose likely benefit is worth whose potential loss? These are questions about society’s values, not about science’s undiscovered secrets. (p. 310)

**Growth of Health Program Evaluation**

Evaluation is a relatively new discipline. Prior to the 1960s, formal, systematic evaluations of social and health programs were conducted rarely, and few professionals performed evaluations as a full-time career (Shadish et al., 1991). With the election of Lyndon Johnson to the presidency in 1964, the United States entered into an era of unprecedented growth in social and health programs for the disadvantaged. Medicare (public health insurance for adults aged 65 and over), Medicaid (public health insurance for low-income individuals), and other health care programs were launched, and Congress often mandated and funded evaluation of their performance (O. W. Anderson, 1985). As public and private funding for evaluation grew, so did the number of professionals and agencies conducting evaluations. Today, evaluation is a well-known, international profession. In many countries, evaluators have established professional associations that hold annual conferences (e.g., American Evaluation Association, Canadian Evaluation Society, African Evaluation Association, and International Organization for Cooperation in Evaluation). Although an association for health program evaluation does not exist in the United States, the American Public Health Association, AcademyHealth (the professional association for health services research), and other groups often serve as national forums for health evaluators to collaborate and disseminate their findings.

Other forces also have contributed to the growth of health program evaluation since the 1960s. Two important factors are scarce resources and accountability. All societies have limited resources to address pressing health problems. In particular, low-income and middle-income countries face the twin problems of severe resource constraints and many competing priorities (Oxman et al., 2010). When resources are scarce, competition for funds can intensify, and decision makers may allocate resources only to programs that can demonstrate good performance at the lowest cost. In such environments, evaluations can provide useful information for managing programs and, if performance is sound, for defending a program’s worth and justifying continued funding. If, however, an evaluation is launched solely to collect information to defend a program’s worth in political battles over resource
allocation, and if the evaluation is conducted in an impartial manner, there is no guarantee that an evaluation will produce results favoring the program.

The trillions of dollars invested globally in health programs has increased the importance of accountability (Oxman et al., 2010). For financial and legal reasons, public and private funding agencies are concerned with holding programs accountable for funds and their disbursement, with an eye toward avoiding inappropriate payments. High-income countries have the greatest expenditures in health programs and, therefore, the greatest potential for waste (Oxman et al., 2010). For performance reasons, funding agencies also want to know if their investments produced expected benefits while avoiding harmful side effects. Similarly, government, employers, and other purchasers of health services are concerned with clinical and fiscal accountability, or evidence that health care systems and providers deliver services of demonstrated effectiveness and quality in an efficient manner (Addicott & Shortell, 2014; Relman, 1988; Rittenhouse et al., 2009; Shortell & Casalino, 2008). Employers also want to know whether their investments in health care improve their employees’ productivity, for example, by collecting information about how quickly workers are back on the job after an episode of care (Moskowitz, 1998).

Another factor stimulating interest in health program evaluation is more emphasis on prevention. Many people and health professionals believe that preventing disease is better than curing it. Because the evidence indicates that much disease is preventable (U.S. Department of Health and Human Services, 2015a), a variety of preventive programs and technologies have emerged to maintain or improve the nation’s health. For example, immunizations to prevent disease, water fluoridation to reduce caries, mammography screening to detect breast cancer, and campaigns promoting the use of bicycle helmets to prevent injuries are common in our society. Healthy People 2000 and its successors, Healthy People 2010 and Healthy People 2020, specify health promotion and disease prevention objectives for the nation and provide a framework for the development and implementation of federal, state, and local programs to meet these objectives (U.S. Department of Health and Human Services, 2015a). As the number of programs has proliferated, so has interest in evaluating their performance in achieving their objectives. However, although prevention and the diagnosis and treatment of illness in its early stages are often advocated because they can save health dollars, preventing illness may either save money or add to health care costs, depending on the intervention and the target population (J. T. Cohen et al., 2008).

Since the 1960s, health program evaluation also has been promoted by public agencies, foundations, and other groups sponsoring a variety of demonstration projects to improve population health and the performance of the health care system or to achieve other goals. Consistent with our nation’s belief in incrementalism in political decision making (Lindblom, 1959, 1979; Marmor, 1998; Shortell & Richardson, 1978), decision makers often desire information about whether a proposed change will work before authorizing changes on a broad scale. To supply this information, decision makers may approve demonstration projects or large-scale social experiments to test the viability of promising solutions to pressing health problems. A prominent example is the
Rand Health Insurance Study, a large-scale experiment in which households were randomly assigned to health insurance plans with different cost-sharing arrangements to determine their impacts on health care utilization and expenditures, health outcomes, and satisfaction with medical and dental care (Aron-Dine et al., 2013; Newhouse & the Insurance Experiment Group, 1993). Because large-scale evaluations are relatively expensive to conduct, controversy may exist about whether their findings are worth the resources invested in them. Nevertheless, decision makers are likely to continue authorizing such projects because they often address critical issues in health policy and because they give decision makers the flexibility to be responsive to a problem while avoiding long-term commitments of resources. Evaluation is an important element of demonstration projects because it provides the evidence for judging their worth.

Another factor contributing to the growth of evaluation is increasing government intervention to fix complex problems in the U.S. health care system. Although the United States expends more per capita for health care than any other country in the world, in 2011 U.S. life expectancy (78.7 years) ranked 26th out of 34 developed countries (Organisation for Economic Co-operation and Development, 2008). Social and physical environments, health behaviors, and genetics account in part for these health patterns, but major problems in the U.S. health care system also contribute to the health deficits. Payment for health care remains largely fee-for-service, which increases overuse and costs, undermining health outcomes and leaving fewer resources for other health-producing societal investments (Evans & Stoddart, 1990). Before 2010, about 18% of Americans were uninsured, and social groups with the worst health were more likely to be uninsured, to have greater unmet needs for preventive and therapeutic care, and were less likely to have doctor visits (Hadley, 2003; National Center for Health Statistics, 2012). For those receiving health care, persistent health care disparities exist across social groups (Agency for Healthcare Research and Quality, 2012), and quality of care is low, with only 55% of Americans receiving recommended care (McGlynn et al., 2003). Patient dissatisfaction with the health care system is high, and health care is often fragmented and provider oriented rather than patient centered (Institute of Medicine, 2001).

To address these and other problems, Congress, with much political rancor, passed the **Patient Protection and Affordable Care Act (ACA)**, the most significant government intervention in the U.S. health care system since the passage of Medicare and Medicaid in the 1960s. Although a key reason for government intervention was to reduce the percentage of uninsured individuals in the United States, another important reason was the escalating costs of health care for federal, state, and local governments, which crowds out resources for other public investments. Between 1989 and 2010, the nation’s health care spending grew from $604 billion to $2.6 trillion, with the public share increasing from 45% in 2010 to a projected 49% in 2022 (Cuckler et al., 2013; Levit et al., 1991). Government intervention in the U.S. health care system likely will be greater in the 21st century than in the 20th, which may stimulate future evaluations of system performance.

Another trend that is increasing health program evaluation is the movement toward **evidence-based practice** in medicine and related fields. One reason for the
low quality of U.S. health care is the lack of evidence about comparative effectiveness, or what services work best, for whom, under what circumstances (Institute of Medicine, 2008; McGlynn et al., 2003). When evidence exists, clinicians may not provide evidence-based care, or the evidence for single conditions is less relevant for patients with multiple chronic conditions, who are the main challenge facing health care systems worldwide (Barnett, 2012; Institute of Medicine, 2001, 2008). To increase medical evidence, Congress created, as part of the ACA, the Patient-Centered Outcomes Research Institute (2015), which has funded over 365 studies of more than $700 million on the comparative effectiveness of alternative clinical services to prevent, diagnose, or treat common medical conditions as well as to improve health care delivery and outcomes. The National Institutes of Health, the Agency for Healthcare Research and Quality, and other federal agencies also fund studies that contribute to the medicine evidence base.

Similar calls to grow the evidence base are voiced in public health practice, management, and other fields (Brownson et al., 2003; Brownson et al., 2009; Walshe & Rundall, 2001). The Centers for Disease Control and Prevention has supported The Community Guide, a website that presents evidence-based recommendations on preventive services, programs, and policies that work or do not work, based on systematic reviews conducted by the Community Preventive Service Task Force. Of equal importance, The Community Guide also indicates gaps in the evidence base in 20 areas where there is insufficient evidence to determine whether an intervention works and where more funding and evaluation are needed to close the gaps (Community Guide, 2015).

A related trend, dissemination and implementation science, also is generating growth in evaluation research. This trend suggests that the supply of evidence is not the problem; rather, the critical issue is that only a fraction of the evidence—about 14%—is translated into clinical practice and that promising results in new publications take 17 years to be implemented widely (Balas & Boren, 2000; L. W. Green et al., 2009; Institute of Medicine, 2013). Implementation science is a relatively new field that identifies the factors, processes, and methods that increase the likelihood of evidence-based (i.e., effective) interventions being adopted and used in medical practice, public health departments, and other settings to sustain improvements in population health (Eccles & Mittman, 2006; Lobb & Colditz, 2013). Implementation science is part of the larger process of translation research, which involves studying and understanding the movement of scientific discovery from “bench-to-bedside-to-population”—or the linear progression from basic science discoveries to efficacy and effectiveness studies, followed by large-scale demonstrations, and, finally, dissemination into practice (Glasgow et al., 2012; Lobb & Colditz, 2013). As dissemination and implementation science continues to grow, so will methods of process and impact evaluation for identifying strategies that can cause greater and quicker uptake of effective interventions into routine use in health organizations.

In the health care system, growth in organizational consolidation and information technology is creating new infrastructures for the evaluation of organizational and health system performance (Cutler & Morton, 2013; Moses et al., 2013). All
sectors of the health care system (insurers, physician offices, hospitals, pharmaceuticals, and biotechnology) are consolidating, primarily horizontally (such as the merger of two or more hospitals) but also vertically (such as the merger of a hospital and a physician group), to lower costs through economies of scale and to gain market power over competitors and other sectors of the U.S. health care economy. The ACA has increased the pace of consolidation and integration by authorizing the development of accountable care organizations, or voluntary groups of integrated delivery systems, hospitals, and other providers that assume responsibility for defined populations of Medicare beneficiaries (Berwick, 2011; Cutler & Morton, 2013; Dafny, 2014).

As the size of health care organizations has increased, so have investments in information technology to lower their costs, increase coordination, and improve the quality and safety of clinical care. Although the evidence is unclear regarding whether information technology has produced these expected benefits, consolidation and information technology have converged to produce massive databases, or “Big Data,” creating new opportunities for advancing evaluation methods, assessing health system performance, and building the evidence base in the foreseeable future (Schneeweiss, 2014). Because the trend toward larger, more complex health care organizations creates management challenges, leaders and administrators are major consumers of evaluation information from Big Data, which can be used to make informed decisions to deliver efficient, effective, equitable care on a population level. Growth in Big Data will require evaluators who are adept in both quantitative and qualitative methods (e.g., the analysis of free-form text in the electronic medical charts of thousands of patients) and who are aware of the strengths and limitations of Big Data (Khoury & Ioannidis, 2014).

As a whole, the forces contributing to the growth of health program evaluation are interrelated and will likely continue well into the 21st century. From these trends, two broad types of health program evaluation have emerged, which are reviewed in the next section.

Types of Health Program Evaluation

Health programs usually are implemented to achieve specific outcomes by performing some type of intervention or service. In general, two basic types of evaluation are conducted in the health field:

- The evaluation of health programs
- The evaluation of health systems

Evaluation of Health Programs

Evaluation of health programs includes programs created to reduce or eliminate a health problem or achieve a specific objective. Healthy People 2020 has established 42 “topic areas” for improving the nation’s health and reducing inequalities, ranging...
from adolescent health and arthritis to tobacco use and vision health (U.S. Department of Health and Human Services, 2015b).

In essence, the topic areas are a comprehensive inventory of the categories of health programs that can be implemented to achieve specific health objectives. For example, water fluoridation programs are implemented to improve oral health, or exercise programs are created to increase physical activity. This type of evaluation assesses the performance of programs developed to achieve health objectives in these and other areas. Some topic areas, such as access to health services and public health infrastructure, overlap with the evaluation of health systems.

**Evaluation of Health Systems**

Aday and colleagues (1998) present a framework for evaluating the performance of health care systems based partly on Donabedian's (1973) earlier work (see Figure 1.1). A health care system has a *structure* defined by federal, state, and local laws and regulations; the availability of personnel, facilities, and other resources; and the organization and financing of care. The structure component also includes the characteristics of the population that the system serves, as well as the physical, social, and economic environments where they live. As a whole, the structure of the system influences the *process* or delivery of health services, which in turn produces *outcomes*, health and well-being. Three criteria are proposed for gauging the worth—or value—of system performance. The three “E’s” define what improvements in health and satisfaction (effectiveness) were produced by health services at what cost (efficiency) and for what population groups (equity). A fourth “E” (ethics) is essential for judging whether a fair or equitable distribution of the costs and outcomes of health services exists among those who need care and those who pay for it. Based on ethical principles of distributive justice, inequitable access to care exists when those who need care the most do not get it. Table 1.1 presents definitions of criteria for assessing effectiveness, efficiency, and equity at the clinical and population levels.

Evaluations of the performance of health care systems typically examine the influence of the structural component on the process of care, or the influence of the structure and process components on the outcomes of care (Begley et al., 2013; Clancy & Eisenberg, 1998; Kane, 1997). For example, an evaluation of the association between the structure and process components of the system was performed by Baicker and associates (2013; see also Finkelstein et al., 2012), who examined Oregon's 2008 expansion of its Medicaid program for low-income adults through a lottery drawing of about 30,000 individuals from a waiting list of almost 90,000 persons. As expected, for persons who met eligibility requirements and enrolled in the program, Medicaid coverage increased the use of health services, but the findings were mixed for quality of care and health outcomes. Medicaid coverage improved rates of diabetes detection and management, improved self-reported health and measures of mental health but not physical health, and reduced financial strain.

The evaluation of health systems also includes the local public health system. Hajat and colleagues (2009) present a framework for evaluating the performance of local public health departments in improving population health and reducing health inequalities.
FIGURE 1.1 • Framework for Evaluating the Performance of Health Care Systems

Source: Figure 1.4 in Evaluating the Healthcare System: Effectiveness, Efficiency, and Equity, 3rd ed., by Lu Ann Aday et al., 2004. Chicago: Health Administration Press.
TABLE 1.1 • Definitions of Effectiveness, Efficiency, and Equity in the Framework for Evaluating the Performance of Health Care Systems

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Level of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clinical</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Clinical effectiveness:</td>
</tr>
<tr>
<td></td>
<td>Improving the health of individual patients through medical care services</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Production efficiency:</td>
</tr>
<tr>
<td></td>
<td>Combining inputs to produce services at lowest cost</td>
</tr>
<tr>
<td>Equity</td>
<td>Procedural equity:</td>
</tr>
<tr>
<td></td>
<td>Maximizing the fairness in the distribution of services across groups</td>
</tr>
</tbody>
</table>

Source: Table 1.1 in Evaluating the Healthcare System: Effectiveness, Efficiency, and Equity, 3rd ed., by Lu Ann Aday et al., 2004. Chicago: Health Administration Press.

(see Figure 1.2). Local health departments (LHDs) are the government entities that are expected to improve population health and reduce health inequalities, particularly in vulnerable social groups, by creating conditions in communities that support good health (Scutchfield & Howard, 2011). LHDs typically have partnerships with other public, private, and voluntary entities, forming a loosely connected, local public health system that coordinates activities to achieve common goals. LHDs operate in a larger environment, or context, that informs their philosophy of public health practice, their mission, and long- and short-term objectives (or “purpose”). Within an organizational structure, LHDs have an infrastructure, or “inputs,” such as personnel, fiscal resources, information, and other resources, which are converted into processes performing the core functions of public health (assessment, policy development, and assurance) and the 10 essential public health services (see Table 1.2), which ultimately drive LHD performance (Hyde & Shortell, 2012). Expected outcomes are improved population health, reduced health inequalities, and a strengthened local public health system.

Public health services and systems research is the name of the relatively new field that applies the methods of health services research—which includes evaluation—to investigate the performance of public health systems (Scutchfield et al., 2009; Scutchfield & Shapiro, 2011). For example, recent studies have examined whether LHDs with greater expenditures per capita have lower mortality and reduced health inequalities. Using
national data, Mays and Smith (2011) report that county-level mortality rates declined 1.1% to 6.9% for each 10% increase in LHD spending. Similarly, Grebowski et al. (2010) examined whether 1990–1997 changes in LHD expenditures per capita were associated inversely with 1990–1997 changes in all-cause mortality rates for Black and White racial groups in U.S. local jurisdictions. Although changes in LHD expenditures were not related to reductions in Black/White inequalities in mortality rates in the total population, inverse associations were detected for adults aged 15–44 and for males. Bekemeier
et al. (2014) also report that LHD expenditures for maternal and child health (MCH) had the expected, inverse relationship with county-level low birth weights, particularly for counties with high concentrations of poverty and for categories of MCH spending based on need.

Historically, the health care system and the public health system have operated largely in isolation in the United States, perhaps because 97% of national health expenditures go to health care services (Institute of Medicine, 2013), and in prior years the health care system had few incentives to build linkages with the public health system. However, by adopting a population perspective—long the hallmark of public health systems—the ACA is creating opportunities for the health care system and the public health system to work together through several mechanisms (Institute of Medicine, 2013). For example, accountable care organizations (ACOs) are responsible for the quality of care in a defined population (Lewis et al., 2013). Fisher et al. (2012) present a framework for evaluating the performance of ACOs, including their impacts on health outcomes, which creates incentives for ACO providers, LHDs, and community partners to deploy population-level strategies to protect the health of the

### TABLE 1.2  The 10 Essential Public Health Services

<table>
<thead>
<tr>
<th></th>
<th>The 10 Essential Public Health Services describe the public health activities that all communities should perform. Public health systems should:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Monitor health status to identify and solve community health problems.</td>
</tr>
<tr>
<td>2.</td>
<td>Diagnose and investigate health problems and health hazards in the community.</td>
</tr>
<tr>
<td>3.</td>
<td>Inform, educate, and empower people about health issues.</td>
</tr>
<tr>
<td>4.</td>
<td>Mobilize community partnerships and action to identify and solve health problems.</td>
</tr>
<tr>
<td>5.</td>
<td>Develop policies and plans that support individual and community health efforts.</td>
</tr>
<tr>
<td>6.</td>
<td>Enforce laws and regulations that protect health and ensure safety.</td>
</tr>
<tr>
<td>7.</td>
<td>Link people to needed personal health services and assure the provision of health care when otherwise unavailable.</td>
</tr>
<tr>
<td>8.</td>
<td>Assure competent public and personal health care workforce.</td>
</tr>
<tr>
<td>9.</td>
<td>Evaluate effectiveness, accessibility, and quality of personal and population-based health services.</td>
</tr>
<tr>
<td>10.</td>
<td>Research for new insights and innovative solutions to health problems.</td>
</tr>
</tbody>
</table>

The framework for the Essential Services was developed by the Core Public Health Functions Steering Committee in 1994. For further information, see the Public Health Functions Project website (http://www.health.gov/phfunctions/).
ACO patient population (Institute of Medicine, 2013, 2014). The ACA’s health insurance exchanges and health information exchanges also are a population approach to health insurance expansion and access to health care (Mays & Scutchfield, 2012; Scutchfield et al., 2012).

Economic evaluation, such as cost-effectiveness analysis and cost-benefit analysis, is an important element of assessing the performance of health programs and health systems. The focus is measurement of the benefits, or outcomes, of a health program or medical technology relative to the costs of producing those benefits. In the face of scarce resources, interventions that produce relatively large benefits at a low cost have greater worth than interventions that offer few benefits and high costs. Specific standards for conducting cost-effectiveness studies have emerged to ensure quality and adherence to fundamental elements of the methodology (Gold et al., 1996). Spurred by the trends toward cost containment and evidence-based medicine and public health practice, the number of cost-effectiveness evaluations in the published literature has skyrocketed since the 1980s. The Tufts-New England Medical Center (2013) Cost-Effectiveness Analysis Registry contains over 10,000 cost-effectiveness ratios for a variety of diseases and treatments. Because new technologies are always being created, cost-effectiveness studies will be a major area of evaluation for many years to come.

This textbook is designed to provide a practical foundation for conducting evaluations in these arenas. The concepts and methods are similar to those found in the evaluation of social programs but have been customized for public health and medical care. Evaluation itself is a process conducted in a political context and composed of interconnected steps; another goal of this book is to help evaluators navigate these steps in health settings. A customized, reality-based treatment of health program evaluation should improve learning and ultimately may produce evaluations that are both practical and useful.

While a key goal of evaluation is discovering whether a health program works, many decision makers ultimately want to know about the generalizability of an evaluation’s findings—that is, do the findings apply to different social groups and settings, variations in the intervention itself, and also for different ways of measuring outcomes (Shadish et al., 2002)? The evaluation methods for examining whether a single health program works cannot answer questions about the generalizability of a program. However, if a sufficient number of evaluations of a program are conducted that address a common evaluation question, and that contain different kinds of social groups and settings with variations in program and outcome, a meta-analysis of their findings may be performed. Meta-analysis is a quantitative technique for synthesizing, or combining, the results from different evaluations on the same topic, which may yield information about whether the findings are robust over variations in persons, settings, programs, and outcomes (Shadish et al., 2002). Although this textbook addresses in depth issues of generalizability of evaluation findings from the evaluation of a single evaluation, the methods of meta-analysis are not covered here but can be found in other references (M. Borenstein et al., 2009; Cooper et al., 2009).
Summary

At its core, evaluation entails making informed judgments about a program’s worth, ultimately to promote social change for the betterment of society. Unprecedented growth in health programs and the health care system since the 1960s is largely responsible for the development of health program evaluation. Other forces contributing to the growth of evaluation include the increasing importance of accountability and scarce resources, a greater emphasis on prevention, more attention being given to evidence-based practice and implementation science, escalating health care costs and government intervention in health systems, organizational consolidation and information technology, and a reliance on demonstration projects. Because many of these trends will continue in the remainder of this century, so will interest in the evaluation of health programs and health systems. To perform either of these two types of program evaluations, an evaluator completes a process composed of well-defined steps. Chapter 2 reviews the elements of the evaluation process.

List of Terms

- Accountability 4
- Demonstration projects 5
- Dissemination 7
- Economic evaluation 14
- Evaluation 2
- Evidence-based practice 6
- Generalizability 14
- Government intervention 6
- Health program 1
- Health systems 2
- Implementation science 7
- Information technology 7
- Meta-analysis 14
- Organizational consolidation 7
- Patient Protection and Affordable Care Act (ACA) 6
- Prevention 5
- Public health services and systems research 11
- Scarce resources 4
- Translation research 7

Study Questions

1. What is the purpose of health program evaluation?
2. What are the two fundamental questions of program evaluation?
3. How is the worth of a health program determined?
4. What are three factors that have contributed to the growth of health program evaluation since the 1960s? Why are they important?
5. What are the major types of health program evaluation, and what are their relationships (if any) to each other?
6. What kinds of evaluations of the health reforms in the ACA might be conducted?