

July 2011

VALUE IN HEALTH CARE

Key Information for Policymakers to Assess Efforts to Improve Quality While Reducing Costs

U.S. Government Accountability Office

GAO 90

YEARS

1921-2011

ACCOUNTABILITY ★ INTEGRITY ★ RELIABILITY

Why GAO Did This Study

The U.S. has devoted an increasing proportion of its economy and federal budget to the provision of health care services, but high levels of spending do not guarantee good care. Policymakers, health practitioners, and others have implemented numerous health care interventions that make discrete changes in the organization of health care services in order to enhance the value of health care—that is, improve the quality of care while reducing costs. Examples include programs to reduce bloodstream infections and to coordinate patient care following hospital discharges.

This report (1) examines the availability of evidence on the effect of selected interventions on quality of care and costs; (2) identifies key dimensions for assessing the strength of such evidence; and (3) examines factors that can facilitate the implementation and replication of health care interventions. GAO identified a broad and diverse set of health care interventions using published and unpublished sources. For 127 of those interventions, GAO analyzed responses to a questionnaire that it sent to persons knowledgeable about available information on the effect of that particular intervention on quality of care and costs. GAO's questionnaire also asked respondents to assess the relative importance of seven factors in the implementation and potential replication of the health care intervention. In addition, GAO consulted the methodological literature and experts on assessing evidence on the effects of health care interventions.

View [GAO-11-445](#) or key components. For more information, contact James Cosgrove at (202) 512-7114 or cosgrovej@gao.gov.

VALUE IN HEALTH CARE

Key Information for Policymakers to Assess Efforts to Improve Quality While Reducing Costs

What GAO Found

About half of the respondents to our questionnaire reported some information on the effect of an intervention on both quality of care and costs—the two types of data needed to determine whether or to what extent a particular intervention enhanced the value of health care. Overall, the vast majority of our respondents reported at least some information on the observed effect of the intervention on quality of care. Relatively fewer—though still over half—of our respondents reported at least some information on the effect of the intervention on costs.

Whether or not policymakers can rely on information that indicates an intervention enhances value depends on the strength of the underlying evidence about quality and cost effects. From studies on the effect of health care interventions on quality of care and costs, policymakers and others can assess the strength and limitations of available evidence along three dimensions. One, the credibility of evidence on the effect of health care interventions on quality of care and costs depends primarily on whether those studies apply rigorous study designs. Two, the applicability of the results of studies to a broader population depends on the extent to which the study population is representative of that larger population. Finally, the capacity of health care interventions for widespread replication can be examined in terms of the consistency of results obtained by each intervention across diverse health care organizational contexts.

Respondents reported, generally by large margins, that leadership support as well as other factors, such as organizational culture and staff resources, significantly facilitated implementation. However, respondents were more divided when asked about the reported effect that health IT had on implementation, and most respondents reported that financial incentives were not a factor in the implementation of the intervention. A majority of respondents reported that each of these factors, with the exception of financial incentives, would be either very or somewhat important if one were to attempt to replicate the intervention as widely as possible.

Progress in achieving greater value in the U.S. health care system will depend, in part, on the availability of information regarding the effect of interventions on quality of care and costs and on how policymakers and others assess and use that information. Information can guide the choices of policymakers among multiple interventions vying for support, but those decisions will have a sounder basis if the information meets certain criteria regarding its content and strength of evidence. At least some information on both cost and quality effects was available for about half of the interventions GAO examined. However, for many interventions the credibility of this information was put into question by widespread reliance on studies that did not incorporate rigorous designs that could isolate the effect of an intervention from other factors.

We requested comments from the Department of Health and Human Services, but none were provided.

Contents

Letter		1
	Background	7
	Respondents Reported Basic Information Needed to Assess Value Available for About Half of Selected Interventions	9
	The Strength of Evidence on the Effect of Interventions Can Be Assessed along Three Dimensions	16
	Leadership Support and Other Factors Reported As Important for Both Implementation and Replication of Interventions	23
	Concluding Observations	28
	Agency Comments	29
Appendix I	Scope and Methodology	30
Appendix II	Types of Health Care Interventions That Seek to Improve the Value of Health Care	39
Appendix III	What Makes Some Study Designs More Rigorous Than Others	41
Appendix IV	Key Questions for Assessing Evidence from Studies of Interventions That Seek to Enhance Value	43
Appendix V	GAO Contact and Staff Acknowledgments	47
Tables		
	Table 1: Type of Information Reported on Quality of Care and Cost Savings for Selected Interventions	10
	Table 2: Frequency of Types of Quality Measures Used to Assess Selected Health Care Interventions	12
	Table 3: Characteristics Distinguishing Rigorous and Weak Study Design Types	18
	Table 4: Reported Effect of Identified Factors on Implementation of Selected Health Care Interventions	24

Table 5: Expected Degree of Importance of Identified Factors for Widespread Replication of Selected Health Care Interventions	27
Table 6: To Assess the Credibility of Attributing Observed Changes in Quality of Care and Costs to the Intervention	44
Table 7: To Assess the Applicability of Study Results for Broader Populations of Interest	45
Table 8: To Assess an Intervention’s Capacity for Widespread Replication	46

Figure

Figure 1: Number of Different Types of Quality Measures Used to Assess the Effects of Selected Interventions	13
--	----

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
CEA	Cost-Effectiveness Analysis
EHC	Effective Health Care Program
EPOC	Effective Practice and Organisation of Care Group
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GDP	gross domestic product
HCIE	Health Care Innovation Exchange
IT	information technology
PPACA	Patient Protection and Affordable Care Act
RCT	randomized controlled trial

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



G A O

Accountability * Integrity * Reliability

United States Government Accountability Office
Washington, DC 20548

July 26, 2011

The Honorable Kent Conrad
Chairman
Committee on the Budget
United States Senate

The Honorable Sheldon Whitehouse
United States Senate

For many years the United States has devoted an increasing proportion of its gross domestic product (GDP) and federal budget to the provision of health care services. National health expenditures rose from 7 percent of GDP—\$308 billion—in 1970 to 16 percent—\$2.2 trillion—in 2008. Over this same period, the proportion of federal budget outlays devoted to health care increased even more rapidly from 10.5 to 32.6 percent.¹ Unless this trend is reversed, spending on health care will consume an escalating share of federal resources, leaving fewer and fewer dollars for other national priorities.

Furthermore, high levels of spending do not guarantee good quality of care. While resources are needed to provide quality care, spending more to increase the number or technical complexity of treatments provided does not always lead to a corresponding

¹Office of the Actuary, Centers for Medicare & Medicaid Services, *National Health Expenditure Tables*, table 1, accessed October 1, 2010, <https://www.cms.gov/NationalHealthExpendData/downloads/tables.pdf>. Total expenditures were adjusted for inflation in 2005 constant dollars. See M Hartman et al, "Health Spending Growth at A Historic Low in 2008," *Health Affairs*, 29:1, (Jan 2010) 148. Total federal budget outlays obtained from U.S. Budget for Fiscal Year 2011, Table 3.2—Outlays By Function and Subfunction: 1962–2015, accessed January 25, 2011, www.gpoaccess.gov/usbudget/fy11/sheets/hist03z2.xls.

increase in the quality of care.² At the same time, studies have documented that many U.S. patients receive care of inconsistent quality as measured in terms of adherence to recognized standards of practice and in terms of clinical outcomes.³

Faced with these challenges, policymakers, health practitioners, and others have looked for ways to enhance the value of our health care by improving the quality of that care and at the same time reducing costs.⁴ To promote greater value across a range of health care settings, they have implemented numerous interventions that make discrete changes in who delivers health care services, how care is organized, or where care is delivered for a specified population. Some of these interventions are designed to restructure the process of health care in ways that guide the behavior of clinicians in fairly defined ways; for example, hospitals have implemented checklists in their intensive care units to help reduce the incidence of bloodstream infections by ensuring more consistent compliance with recommended procedures. Other interventions focus on restructuring how providers are paid to encourage them to produce greater value without trying to specify what steps they should take to achieve that objective. For example, payments have been structured so that providers earn more if their patients experience high quality care while overall costs are held in

²E.S. Fisher and H.G. Welch, "Avoiding the Unintended Consequences of Growth in Medical Care: How Might More Be Worse?" *Journal of the American Medical Association*, vol. 281, no. 5 (1999): 446-453; E.S. Fisher, et al., "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care," *Annals of Internal Medicine*, vol. 138, no. 4 (2003): 273-287; E.S. Fisher, et al., "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care," *Annals of Internal Medicine*, vol. 138, no. 4 (2003): 288-298; and Joseph P. Newhouse and the Insurance Experiment Group, *Free for All? Lessons from the RAND Health Insurance Experiment* (Cambridge, Mass.: Harvard University Press, 1993).

³For example, a seminal study found that on average Americans receive about half of recommended medical care processes, and the Institute of Medicine reported that at least 1.5 million preventable adverse drug events injure patients in the U.S. each year. See Elizabeth A. McGlynn, et al., "The Quality of Health Care Delivered to Adults in the United States," *The New England Journal of Medicine*, 348:26 (June 26, 2003), 2643, and Institute of Medicine, *Preventing Medication Errors*, (2006), 3.

⁴In this report we define value in health care as lowering or holding costs constant while sustaining or increasing quality of care.

check. Still other interventions are designed to motivate or assist patients to take actions that will enhance their own health and thereby reduce their need for health care services, ranging from self-management of chronic conditions such as diabetes or heart failure to maintenance of recommended drug regimens.⁵ The 2010 Patient Protection and Affordable Care Act (PPACA) includes multiple provisions that support one or more of these approaches to enhance the value of health care.⁶

The wide array of different interventions gives policymakers the opportunity to help improve the value of U.S. health care by supporting those interventions for which there is good evidence that they improve quality and reduce costs. To identify these interventions, policymakers need information on the effects of interventions on both quality of care and costs, and they need that information to be credible. It is therefore important for policymakers to be able to weigh the strengths and limitations of the evidence that an intervention, on net, has led to positive changes in quality of care and costs. Finally, interventions may vary in their potential for replication; that is, their effects on quality of care and costs may differ substantially among the organizations—such as hospitals and physician practices—that attempt to implement them. Therefore, assessments of the evidence for different health care interventions will be more complete if policymakers consider the interventions' effects on quality of care and costs across different contexts. Policymakers may also find it useful to know what factors may inhibit or facilitate the implementation and replication of interventions across varied organizational contexts.

To assist policymakers in their efforts to enhance the value of health care, you requested that we provide you with information about health care interventions. This report (1) examines the

⁵See appendix II for a more complete description of different types of health care interventions that are intended to enhance value.

⁶Pub. L. No. 111-148, 124 Stat. 119 (2010). Prominent examples include the establishment of a Center for Medicare and Medicaid Innovation within the Centers for Medicare & Medicaid Services (sections 3021 and 10306), the Medicare Shared Savings Program (sections 3022 and 10307), and the Hospital Readmissions Reduction Program (sections 3025 and 10309).

availability of information on the effect of selected health care interventions on quality of care and costs; (2) identifies key dimensions for assessing the strengths and limitations of available evidence on the effect of interventions on quality of care and costs; and (3) examines factors that can facilitate the implementation and replication of health care interventions.

To examine the availability of information on the effect of selected health care interventions on quality of care and costs, we drew on multiple sources to identify a broad and diverse set of interventions that related to value in health care. In addition to an extensive literature review, we conducted a comparable examination of other, nonbibliographic sources including a database on quality improvement initiatives maintained by the Agency for Healthcare Research and Quality (AHRQ) and materials from presentations at research conferences. The latter two sources allowed us to include interventions that had not yet been described in academic or professional literature. From these various sources we selected for review a set of 239 interventions that appeared to meet all of the following criteria: (1) the intervention had implemented a discrete change in the organizational structure or process of health care delivery; (2) available descriptions of the intervention suggested that it both improved quality and reduced health care costs (or held one constant while improving the other); and (3) the intervention addressed issues relevant to the U.S. health care system.⁷ (See app. I for a more complete description of the sources and methods we used to identify interventions.)

Because the documentation that we obtained on these interventions varied widely in focus, substantive content, and date of issue, we chose to collect detailed information from individuals with expert knowledge of each intervention using a standardized data collection instrument. Thus, we developed a Web-based questionnaire that we sent to researchers who were primary authors of articles identified in our literature search or prepared other materials describing these interventions and their results. The

⁷This criterion excluded interventions that focused on diseases or conditions that rarely occur in the U.S. or that addressed the particular challenges of delivering health care services in lesser-developed countries.

questionnaire included a mix of open-ended and closed-ended questions that examined what information was available on the effect of the intervention on quality of care and costs as well as what factors had facilitated or impeded implementation and replication of the intervention. With respect to quality of care, we asked respondents to describe up to five key measures that were used to assess the effect of their intervention on quality of care and the magnitude of change observed in those measures relative to a baseline or a control group. Regarding costs, we asked respondents to report the amount of any savings attributable to the intervention as well as the methods and information used to calculate those savings.

To identify key dimensions for assessing the strengths and limitations of available evidence on the effect of interventions on quality of care and costs, we reviewed the relevant methodological literature on conducting systematic reviews and evaluations of health care interventions. We also consulted with several subject matter experts and obtained their reaction to the set of criteria that we identified through that review to help policymakers critically assess the information presented to them on value-enhancing interventions. It was beyond the scope of this engagement to apply this set of criteria to individual interventions. However, drawing on the documentation collected in identifying interventions related to value, we were able to categorize the types of study designs employed by studies that examined the interventions for which we received responses to our questionnaire.

With respect to factors that facilitated or impeded implementation and replication of the intervention, we identified from relevant literature seven factors that have been found to facilitate or impede efforts to change the organizational structure and process of health care delivery: leadership support, organizational culture, staff resources, health information technology (IT), availability of tools and activities to standardize care, financial resources, and financial incentives. We asked our questionnaire respondents to assess the extent to which each of these seven factors facilitated or impeded the implementation of their intervention and the expected degree of

importance that each of these factors could have on attempts to replicate the intervention to the widest scale possible.⁸ To gain contextual understanding of how these factors affect implementation and replication, we asked respondents to provide a narrative description of how each factor facilitated or impeded implementation of the intervention and why these factors would be important for widespread replication, respectively. (See app. I for a more extensive description of the information collected through our questionnaire.)

We received usable responses for 127 of the 239 selected interventions. These 127 interventions applied a broad range of value-enhancing strategies in different health care settings and among diverse patient populations, including enhanced management of patients with chronic conditions and coordination of care across multiple providers (see app. II for a complete list of the different types of interventions included). Although our efforts to identify relevant interventions were extensive, we could not ensure that every intervention meeting our selection criteria had been identified. As a result, our findings are limited in scope to the interventions for which we received completed questionnaires and cannot be generalized to all value-enhancing health care interventions.

We conducted this performance audit from April 2009 to July 2011, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

⁸Respondents had the opportunity to provide the same information for any additional factors, beyond the seven listed on the questionnaire, that they had found also affected implementation or replication of the intervention. However, only a few respondents identified any additional factors.

Background

The methodological literature provides insight into conducting systematic assessments of evidence for health care interventions that change the delivery or structure of care. Furthermore, the literature on organizational change is pertinent to understanding the key factors that can facilitate or impede implementation and replication of such health care interventions.

Systematic Assessments

Applied social science research has developed a core set of methodological questions and approaches for assessing the effect of programs or other interventions on a wide range of organized activities.⁹ They address two key issues: how best to determine the independent effect of a program or intervention and how best to generalize from the results obtained from one or more studies to broader populations of interest.¹⁰ A number of organizations have developed more specialized guidance for applying these general methodological principles to health care interventions. For example, the Effective Practice and Organisation of Care Group (EPOC) is a component of the Cochrane Collaboration—an international network of individuals who analyze the effect of health care interventions—which focuses on interventions that change the practice of care and the delivery of health care services. EPOC provides guidance to researchers on how best to prepare systematic reviews of such interventions in order to synthesize the information available in multiple studies.¹¹ AHRQ's Effective Health Care Program (EHC) and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group have developed similar guidance, though both of these efforts focus more on assessing alternative medical treatments rather than alternative approaches for organizing health care services.

⁹See WR Shadish, TD Cook, and DT Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, (Boston: Houghton Mifflin Co., 2002) for an authoritative explanation of these questions and approaches.

¹⁰In the methodological literature, these issues are usually referred to as internal validity and external validity.

¹¹This guidance is available online at <http://epoc.cochrane.org/epoc-resources-review-authors>.

Key Factors That Contribute to Organizational Change

Research on organizational change has identified certain factors as key contributors to successful implementation of health care interventions for quality improvement.¹² For example, the literature consistently cites leadership support as essential for successful implementation of quality-of-care interventions within health care settings. Further, it describes the role of leaders in promoting the adoption of interventions by their organizations, in winning acceptance among affected staff members for the changes those interventions entail, and in marshalling sufficient resources for the intervention to succeed. In addition, this literature has shown that organizations vary in their attitudes, beliefs, and values, and that this “organizational culture” can either promote or inhibit change. Organizations tend to achieve quality of care improvement more readily if they have a culture with such characteristics as receptiveness to change, placing high value on ensuring the quality of care provided, and prizing innovation with a willingness to take risks. The literature also cites the role of infrastructure factors, such as the sufficiency and appropriateness of staff resources and the adequacy of existing health information technology (health IT) systems, in the successful implementation of quality improvement interventions. Another factor cited in the literature is the availability of previously developed tools and procedures for standardizing health care processes—such as checklists or guidelines—as well as other types of technical assistance that can facilitate the implementation of a given intervention. Additionally, the literature has pointed to financial factors that affect the implementation of interventions for quality improvement, including both the level of financial resources needed to sustain an intervention and the use of financial incentives to promote quality enhancement activities. Financial incentives represent a particular application of financial

¹²See, for example, J. Øvretveit. *Does improving quality save money? A review of evidence of which improvements to quality reduce costs to health service providers.* (London: The Health Foundation, September 2009); C. Homer and R. Baron. “How to Scale Up Primary Care Transformation: What We Know and What We Need to Know?” *Journal of General Internal Medicine*, vol. 25, no. 6 (2010); and E. Ferlie & S. Shortell. “Improving the Quality of Health Care in the United Kingdom and the United States: A Framework for Change.” *The Milbank Quarterly* vol. 79, no. 2, 2001. For further description of research studies that identify factors that can affect the implementation and replication of health care interventions, see appendix I.

resources that involve the contractual or other provisions that determine how much health care providers are paid and for what. Such financial incentives affect who benefits from and who pays for the cost of an intervention. This in turn can facilitate or impede the implementation and replication of interventions.

**Respondents
Reported Basic
Information Needed
to Assess Value
Available for About
Half of Selected
Interventions**

About half of the respondents to our questionnaire reported basic information on the effect of their intervention on both quality of care and costs—the two types of data needed to determine whether or to what extent a particular intervention enhanced the value of health care. Overall, the vast majority of our respondents reported at least some information on the observed effect of their intervention on quality of care. Relatively fewer—though still over half—of our respondents reported at least some information on the effect of their intervention on costs.

**About Half of Respondents
Reported Basic
Information on the Effect
of Selected Interventions
on Both Quality of Care
and Costs**

The ability of policymakers to identify interventions that substantially improve quality and reduce costs depends on the availability of basic information on the size of the effect of an intervention on both quality of care and costs. These are the two types of data needed to determine whether or to what extent a particular intervention enhanced the value of health care. Just over half of the respondents to our questionnaire reported such basic information on their interventions. Sixty-four of 127 respondents reported information on both improvements observed in at least one quality measure and a specific amount of cost savings (see table 1). For the remaining interventions, the missing information most often concerned the effect of the intervention on costs. Furthermore, even fewer respondents, 45, reported improvements observed in at least one quality measure and a specific amount of cost savings that accounted for the costs of implementing their intervention—net cost savings.

Table 1: Type of Information Reported on Quality of Care and Cost Savings for Selected Interventions

Type of information reported	Number of Interventions
Reported observed improvements in at least one quality measure	114
Reported a specific amount of cost savings	72
Reported observed improvements in at least one quality measure and a specific amount of cost savings	64
Reported observed improvements in at least one quality measure and a specific amount of net cost savings	45

Source: GAO analysis of responses to GAO questionnaire.

Note: Based on responses for all 127 interventions in our analysis. Respondents who reported a specific amount of net cost savings stated that this amount of reported savings took account of the costs of implementing their intervention.

Information on the Effect of Selected Interventions on Quality of Care Was Frequently Reported

Compared to information on both quality of care and costs, information on the effect of selected interventions on quality of care alone was more frequently reported. The vast majority of respondents to our questionnaire reported at least some information on the observed effect of their intervention on quality of care. Specifically, 114 of 127 respondents reported improvements in one or more measures used to assess the effect of their intervention on various aspects of care quality.¹³ Of these, 112 respondents reported a specific magnitude of improvement observed in at least one quality measure in terms of a percentage change or other quantitative measurement.¹⁴ Additionally, 2 respondents reported improvement in at least one quality measure, but did not report a specific magnitude of improvement. In contrast, the remaining 13 respondents did not report sufficient information to determine whether their intervention had any effect on quality of care. Six of 127 respondents described one or more measures used to assess the effect of their intervention on different aspects of care quality, but did not report a magnitude of

¹³ Respondents listed up to five measures used to assess the effect of their intervention on different aspects of care quality; we did not attempt to verify the appropriateness of the quality measures selected.

¹⁴ Respondents reported a magnitude of change observed in quality measures attributable to their intervention relative to a baseline assessment prior to implementation of the intervention or a control group that did not experience the intervention. We did not assess the extent to which observed changes in quality measures were due to the intervention and not to other factors.

improvement observed in these measures. Seven respondents did not report any information on the measures used to assess the effect of their intervention on aspects of care quality.

Respondents reported that the effect of their intervention on quality of care was assessed using a range of measures that generally fell into five broad types reflecting different aspects of care quality (see table 2).¹⁵ Respondents most frequently described one or more quality measures that were used to assess the effect of their intervention on outcomes resulting from care. Specifically, 82 respondents reported that the effect of their intervention on quality of care was assessed using outcome measures such as patient mortality, the overall physical and emotional health of a patient, or the level of stress reported by a patient caregiver.¹⁶ In addition, 56 respondents described one or more measures that assessed the effect of their intervention on the amount of health care services consumed.¹⁷ These measures included the length of hospital stay, the number of emergency department visits, and the number of hospital readmissions for a specified population. Forty-four respondents described measures that assessed the effect of their intervention on processes of care. Process-of-care measures assess the extent to which the care provided to a patient was appropriate based on current professional knowledge and the particular circumstances.¹⁸ For example, process-of-care measures could examine whether diabetes patients had received foot exams, eye exams, and regular glucose monitoring at specified intervals.

¹⁵We categorized quality measures into types that are largely based on the measure domains laid out by the Agency of Healthcare Research and Quality (AHRQ) in its National Quality Measures Clearinghouse. See appendix I for more information.

¹⁶A caregiver is a family member, friend, or other individual who is responsible for meeting the physical and psychological needs of a patient. Caregivers are distinct from clinical providers such as physicians or nurses. Caregivers are the targeted population of some interventions.

¹⁷AHRQ includes use of health care services measures as one of its measure domains, but it specifies that these measures should be used to monitor trends and are not direct measures of the quality of clinical care provided by health care professionals or organizations.

¹⁸Many process of care measures are based on scientific evidence or established guidelines developed by professional organizations.

Fewer respondents described measures that assessed quality in terms of the experience of a patient or caregiver or the structure of care.¹⁹

Table 2: Frequency of Types of Quality Measures Used to Assess Selected Health Care Interventions

Type of measure	Examples of specific quality measures reported	Number of interventions
Outcome	<ul style="list-style-type: none"> Mortality Quality of life Prevalence of hospital-acquired pressure ulcers Caregiver strain index or stress level 	82
Use of services ^a	<ul style="list-style-type: none"> Proportion of patients (participating in the intervention) admitted to the hospital Length of hospital or ICU stay Duration of mechanical ventilation use 	56
Process	<ul style="list-style-type: none"> Diabetes quality indicators such as number of patients who received a foot exam, eye exam, and glucose monitoring Patient adherence to prescribed medications 	44
Experience	<ul style="list-style-type: none"> Patient satisfaction with care Patient and caregiver feeling of support or connection with provider 	31
Structure	<ul style="list-style-type: none"> Physician, nurse, or other clinician satisfaction Physician knowledge of local guidelines of antibiotic prescribing practices Staff injury 	27

Source: GAO analysis of responses to GAO questionnaire.

Note: One hundred twenty of 127 respondents to our questionnaire reported one or more measures used to assess the effect of their intervention on different aspects of care quality. We categorized these quality measures into five broad types based on measure domains in AHRQ's National Quality Measures Clearinghouse. We excluded from this analysis eight measures reported by six respondents who did not clearly specify what aspect of care quality was assessed by those measures. Most respondents reported that the effect of their intervention was assessed using more than one type of quality measure.

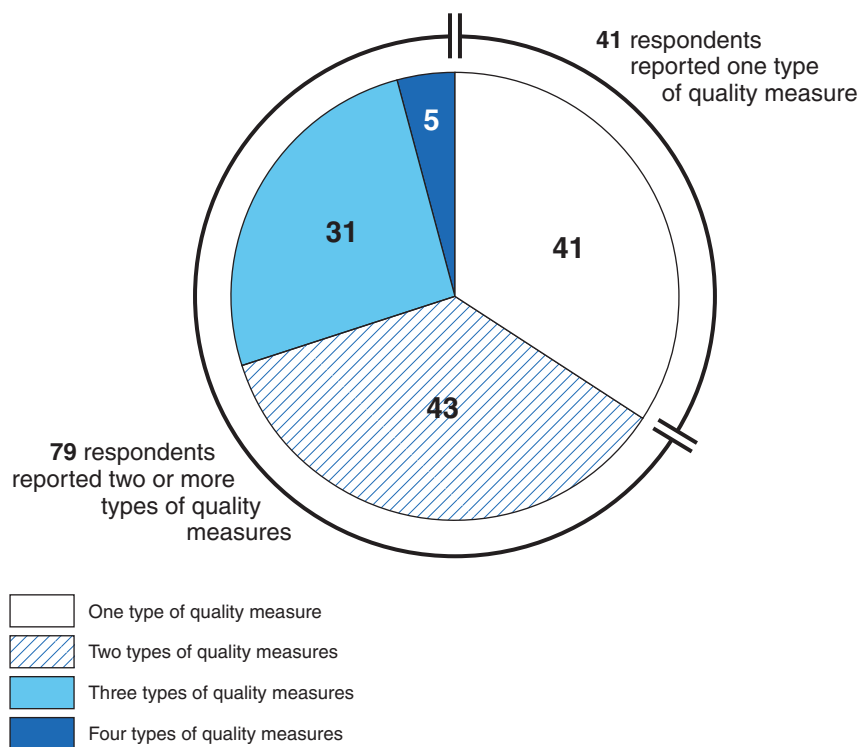
^aAHRQ does not consider use of services measures to be direct measures of the quality of clinical care.

Although the information provided by any one type of quality measure is limited, most of our respondents reported that the effect of their intervention on quality of care was assessed using more than one type of quality measure. Each type of quality measure

¹⁹Experience of care measures assess quality of care in terms of the perspective of patients and caregivers toward the care that they participate in. These measures include patient and caregiver satisfaction. Structure of care measures assess a feature of a health care organization or clinician relevant to their capacity to provide health care. These measures include physician satisfaction and physician knowledge.

offers insight into a particular domain of quality such as outcomes of care, processes of care, or experience of care. Just 41 respondents reported that only one type of measure was used to assess the effect of their intervention on quality of care (see fig. 1). For most—79—of the interventions in our review, respondents reported that the effect of their intervention on quality of care was assessed using measures belonging to two or more different types of quality measures, thereby providing a broader perspective on the effect of the intervention on quality of care.

Figure 1: Number of Different Types of Quality Measures Used to Assess the Effects of Selected Interventions



Source: GAO analysis of responses to GAO questionnaire.

Note: One hundred twenty of 127 respondents to our questionnaire reported that the effect of their intervention on quality of care was assessed using one or more of five different types of quality measures: outcome, use of services, process, experience, and structure. Another 7 respondents did not describe any measures used to assess the effect of their intervention on quality of care.

Information on the Effect
of Selected Interventions
on Costs Was Less
Frequently Reported

Somewhat fewer respondents to our questionnaire reported information on the effect of their interventions on costs than quality of care. Specifically, 72 of 127 respondents reported a specific amount of change in costs—cost savings.²⁰ Respondents most frequently reported that costs were assessed by calculating the total dollars saved or the average dollars saved per person annually. Respondents less frequently reported that costs were assessed by calculating the financial return on investment, percentage change in total health care costs per patient, or an alternative cost metric such as dollars saved per member per month for patients participating in a certain health care plan.²¹ In contrast, the remaining 55 respondents did not report sufficient information to determine whether their intervention had any effect on costs. Nine of 127 respondents reported that costs were assessed, but did not report a specific amount of cost savings. Forty-five respondents reported that cost savings were not assessed, and one respondent did not report any information on whether cost savings were assessed.

Most, but not all, of the respondents who reported a specific amount of cost savings stated that these cost savings accounted for the costs associated with implementing the intervention. Among the 72 respondents who reported a specific amount of cost savings, 51 respondents reported net cost savings that took account of implementation costs; another 20 respondents reported gross cost savings that did not take implementation costs into account.²² When asked to provide additional detail on their implementation cost calculations, 35 respondents reported that the cost savings took account of both start-up costs associated with developing and initially implementing the intervention as well as ongoing costs

²⁰Respondents who reported that cost savings attributable to their intervention were assessed also reported information on the specific amount of costs saving annually and how these cost savings were calculated. We did not assess the extent to which these cost savings were due to the intervention and not to other factors.

²¹The financial return on investment is the amount of money gained or lost as a result of an intervention relative to the money invested in the intervention.

²²One respondent who reported specific cost savings did not respond when asked if the savings were net or gross.

associated with operating and maintaining the intervention over time. Two respondents reported that cost savings took account of start-up costs but not ongoing costs to maintain the intervention, and 19 reported that cost savings took account of ongoing costs but not start-up costs.²³

The interventions we reviewed also varied in the extent to which the reported cost savings attributed to them were based on information directly related to the intervention. Forty-nine respondents reported that cost savings were calculated using only data that were collected specifically to assess the effect of their intervention on costs. In contrast, 26 respondents reported that cost savings were calculated using a mix of data that were collected specifically to assess the intervention and data from a secondary source such as published literature or a national database. For example, one respondent reported cost savings attributable to an intervention designed to improve patient self-management of asthma based on data that were collected on changes over time in the actual number of health care encounters for patients enrolled in the program and the estimated costs for those encounters derived from national averages for several types of health care services such as hospital days or emergency department visits. While data from secondary data sources may provide otherwise missing information needed to estimate the cost savings achieved by an intervention, the relevance of such secondary data to that particular intervention may be open to question, which makes the accuracy of the cost savings estimate more uncertain.

²³Some respondents who reported net cost savings also took account of other implementation costs such as costs associated with developing and implementing a study of the intervention. One respondent who reported net cost savings did not specify what costs were included in the costs of implementing the intervention.

The Strength of Evidence on the Effect of Interventions Can Be Assessed along Three Dimensions

Policymakers and others can assess the strength and limitations of available evidence from studies on the effect of health care interventions on quality of care and costs along three dimensions. One, the credibility of evidence on the effect of health care interventions on quality of care and costs depends primarily on whether those studies apply rigorous study designs. Two, the applicability of the results of studies to a broader population depends on the extent to which the study population is representative of that larger population. Finally, the capacity of health care interventions for widespread replication can be examined in terms of the consistency of results obtained by each intervention across diverse organizations. Appendix III provides a more detailed explanation of what makes some study-design types more rigorous than others and appendix IV presents a list of key questions that describe the information that policymakers can look for to assess the evidence provided by particular studies along these three dimensions.

Examining the Credibility of Evidence on the Effects of Health Care Interventions

For policymakers and others, the benefit obtained from basic information on the effect of interventions on quality of care and costs depends in large part on the strength of that evidence. Information based on weak evidence can provide policymakers a misleading indication of an intervention's potential to enhance value. For example, the direction and magnitude of the changes in quality of care and cost reported for the 127 interventions examined through our questionnaire could deviate substantially from the actual impact of those interventions, depending on the characteristics of the studies that generated that reported information. To determine what information has the kind of evidentiary support that they can rely on, policymakers can assess the strengths and limitations of studies that examine health care interventions of interest along three broad dimensions.

The first of these dimensions is the credibility of evidence that attributes any changes in quality of care and costs to those interventions. The methodological experts we consulted uniformly emphasized the primacy of study design in determining the credibility of evidence on the effect of health care interventions on quality of care and costs. Observed changes in quality of care and costs that one might attribute to a health care intervention may in

Credible Evidence Depends on Rigorous Study Designs

- A 2009 study assessing a surgical safety checklist using a simple pre/post design found the rate of surgical complications declined 36 percent after the checklist was implemented.
- Critics noted that the pre/post design did not control for potential confounding factors, such as other quality improvement initiatives occurring at the same time.
- A 2010 study of a different surgical safety checklist used a controlled before and after design that compared a group of hospitals that implemented the checklist to a separate control group of similar hospitals that did not implement the checklist.
- Surgical complication rates in the control group hospitals did not change significantly, while they decreased 39 percent in the hospitals that implemented the checklist. That contrast provides a substantial basis for crediting the checklist for the decrease in complications as opposed to any other factors.

fact be due in large part to the effect of a wide variety of other factors. The choice of study design type is critical because rigorous designs have the capacity to isolate the effects of a health care intervention from other factors that may affect changes in quality of care and costs.

The methodological literature we reviewed identifies several different study design types that have sufficient rigor to isolate the effect of interventions on quality of care and costs. They include randomized controlled trials (RCTs), interrupted time series studies, and controlled before and after studies.²⁴ RCTs and controlled before and after studies both use control groups—consisting of study participants who are not exposed to the intervention—to adjust for the effect of other factors besides the intervention.²⁵ Interrupted times series studies do not use control groups; instead they rely on analyzing data collected at multiple time points both before and after an intervention is implemented to adjust for other factors. (See app. III for more information on how these study design types isolate the effect of an intervention.)

In contrast, according to the methodological literature we reviewed, some other types of study designs lack the capacity to isolate the effect of a health care intervention from that of other factors. For example, a simple pre/post study that assesses quality of care and costs once before an intervention is implemented and a second time after implementation of the intervention has no mechanism analogous to a control group to take account of the effect of other factors. The same is true for post-only studies that rely entirely on data collected after an intervention was implemented. With studies using these types of designs, there is no way to determine how much of the difference observed between the pre and post measurements, or among any groups following an intervention, was

²⁴These three design types can be used in a wide variety of circumstances. There are other rigorous design types that apply in more specialized circumstances, such as regression discontinuity designs.

²⁵In an RCT the study population is allocated randomly between control and treatment groups while in controlled before and after studies the treatment and control groups are formed on some nonrandom basis such as self selection or judgments made by clinicians or program administrators.

due to the intervention and not to other factors. Consequently, such studies will not provide policymakers credible information about the extent to which the intervention itself affected both quality of care and costs.

Table 3 describes key distinguishing characteristics to help policymakers identify the type of study design employed in a study of an intervention.²⁶

Table 3: Characteristics Distinguishing Rigorous and Weak Study Design Types

Relative design strength	Design type	Distinguishing characteristics
Rigorous	Randomized controlled trial	Outcomes compared for study participants randomly allocated to intervention and control groups
	Controlled before and after study	Outcomes measured before and after implementation of intervention in nonrandomly selected intervention and control groups
	Interrupted time series study	Trends in measured outcomes examined over many time points both before and after intervention implemented
Weak	Pre/post study	Outcomes measured only for those exposed to an intervention, once (or a few times) before and once (or a few times) after implementation of the intervention.
	Post-only study	Outcomes measured only after the intervention was implemented, with or without a control group

Source: GAO synthesis of methodological literature

Note: The three rigorous design types can be used in a wide variety of circumstances; there are other rigorous design types that apply in more specialized circumstances.

Among studies addressing the effect of health care interventions on quality of care and costs, a range of rigorous to weak design types are used. For example, among the 127 interventions for which we received responses to our questionnaire, we found 22 interventions with studies involving RCTs and another 11 interventions assessed using controlled before and after studies. However, for a substantially larger number of the 127 interventions, the studies we identified employed the types of study designs that do not isolate

²⁶The methodology section of a study report may name the type of study design used or provide the information needed to identify the design type from its distinguishing characteristics.

the effect of the intervention from other factors. Specifically, the results for 67 interventions were based on pre/post studies, and another 19 were based on post-only studies of one kind or another. In this one, diverse set of interventions that we reviewed, policymakers could find credible evidence based on rigorous study designs concerning the effects of certain interventions on quality of care and costs; however, for many other interventions such studies were lacking.

In addition to study design, the methodological literature we reviewed emphasized the importance of how a study is conducted. Even rigorous study designs can lose their capacity to isolate the effect of an intervention on quality of care and costs if researchers do not adhere to the requirements of those designs. Thus, assessments of the strengths of study results should consider how well the study design was implemented.

One component of a study's implementation that policymakers can examine involves the selection and management of control groups used in the study. In order to isolate the effects of an intervention, the control group has to be equivalent to the treatment group—except for the latter's exposure to the intervention. According to the methodological literature we reviewed, that equivalence can be compromised in a number of ways. In the case of RCTs, for example, allocation to treatment and control groups may not be truly random if there are flaws in the process for assigning study subjects to those groups. Moreover, for both RCTs and controlled before and after studies, losing a disproportionate number of study participants from either treatment or control groups can also undermine their equivalence.²⁷

Another component of a study's implementation that policymakers can examine concerns the measures and procedures adopted for data collection. According to the methodological literature we reviewed, a study will produce stronger evidence when it employs

²⁷If, for example, a substantially larger proportion of control group members drop out of a study than do members of the corresponding treatment group, the results of the study can be biased by any differences that distinguish those study participants that are lost from those that remain in the study and provide the data that generate the study's results.

measures that are recognized as valid and reliable.²⁸ For example, central line-associated bloodstream infections can be tracked using a surveillance measure developed by the Centers for Disease Control (CDC) or with less labor-intensive measures that draw on administrative data. Clinicians consider the CDC measure to be the most valid and reliable measure for this type of infection because it calls for laboratory confirmation of identified infections and it accounts for varying risks of infection based on the number of days that a central line catheter is in place.²⁹ In addition, the data for those measures should be collected at the same time and in the same way from all groups in the study. Any systematic inconsistencies in how data are collected for a study can skew the results.³⁰

Examining the Applicability of Study Results to Broader Populations

If a study produces credible evidence that a health care intervention has a positive effect on both quality of care and costs within the population it examined, a second dimension that policymakers and others can assess concerns the scope of that effect—for what broader populations or groups are the results applicable? Applicability depends on the representativeness of the study population for a broader population of interest. The methodological literature identifies two different approaches for establishing representativeness: (1) randomly selecting the study population from a known universe, or (2) examining the degree to which a study population matches a given broader population on characteristics relevant to the intervention. The first approach, random selection, intrinsically makes the study population

²⁸Study reports may cite previous research to demonstrate that the study used measures with established validity and reliability.

²⁹P.J. Pronovost et al, "Preventing Bloodstream Infections: A Measurable National Success Story in Quality Improvement," *Health Affairs*, 30:4 (April 2011), pp. 629-30.

³⁰For example, if those collecting data on patient outcomes know which study participants are in the treatment group and which are in the control group, that knowledge can lead them to assess members of the treatment group differently. One technique for addressing this issue is blinded assessment of outcomes, which conceals from those making these assessments knowledge of which study participants are in the treatment and control groups.

Applicability of a Study Population with Nontypical Characteristics

- The Medicare Physician Group Practice (MPGP) Demonstration provided selected group practices bonus payments for meeting quality and cost constraint targets.
- Physician practices were selected for the demonstration based in part on their having sophisticated information systems to track patient services and outcomes.
- All 10 MPGP practices were very large, with more than 200 physicians. However, fewer than 5 percent of all U.S. physicians report practicing in groups of 50 or more.
- An evaluation of the MPGP demonstration noted that its modestly positive outcomes were applicable to physicians with similar characteristics, i.e., those who belong to very large practices. Whether or not physicians belonging to smaller group practices would respond similarly to the same set of incentives cannot be determined from the MPGP demonstration results.

representative of the particular universe from which it was selected and the study results applicable to that population.³¹

The second approach for establishing representativeness—examining the extent of similarity between the study population and a broader population of interest—can be used by policymakers whenever the study population was not chosen randomly or the broader population of interest to policymakers is not the universe from which the study population was selected. Policymakers can assess the degree of similarity between the study population and a broader population through an examination that focuses on two issues: (1) identifying characteristics where the study population and broader population of interest differ and (2) assessing whether any differences found could influence the effect of the intervention on quality of care and costs (see app. IV).³²

Major differences between a nonrandomly selected study population and a broader population of interest to policymakers should raise questions about the applicability of the study's results for that broader population. For example, an intervention to improve care coordination for patients with diabetes might be implemented and assessed in a few academic medical centers. In that situation, the representativeness of the study population for all patients with diabetes could come into question on at least two counts—the kind of care provided in an academic medical center might well differ from that usually provided by community-based providers and the patients treated by academic medical centers might have a higher level of severity than diabetics treated elsewhere. If patients in the study received a different overall set of services, that could affect the impact of the intervention on those patients even if the intervention itself were implemented the same way for the two populations. Similarly, an intervention could have a more pronounced effect on patients with a higher level of severity, or the

³¹Study results based on a random sample will have some uncertainty due to sampling error. However, because sampling error is random and therefore unbiased, it can be quantified in terms of a confidence interval and the study results are representative of the universe from which the sample was selected.

³²References to such analyses, if they are conducted, may appear in the discussion or comment sections of study reports.

intervention might work less well for such patients. Thus, to establish the applicability of the study results to a broader population of diabetic patients, studies of the intervention would need to provide evidence that the differences between the study population and the broader population of diabetics would not affect the performance of the intervention.

Examining the Capacity of Interventions for Widespread Replication

A third dimension on which policymakers and others can assess the strength of evidence for health care interventions concerns the capacity of an intervention for replication across diverse organizations. Because organizations vary across the factors that affect the implementation of health care interventions, including leadership, organizational culture, and staff and financial resources, a particular intervention may work more or less well depending on the organizational environment in which it operates. As a result, some organizations may be more receptive to a particular value-enhancing intervention than others. That, in turn, can make it more difficult to take an intervention that proved successful in a small number of organizations and replicate it widely across many others. However, some interventions have produced positive results on quality of care and costs in a range of different organizations, which suggests that they may be less sensitive to varying circumstances across organizations.

Inconsistent Results across Diverse Sites Indicates Restricted Capacity for Replication

- An intervention to facilitate prevention of pressure ulcers enrolled 21 nursing homes in 4 states during 2006/2007.
- All 21 nursing homes implemented the intervention's core elements—streamlining patient documentation forms with integration into an IT reporting system to promote accurate and timely monitoring of patient condition.
- In aggregate, participating nursing homes achieved a 9 percent improvement in CMS's quality measure for preventing pressure ulcers among high risk patients.
- A subset of "high implementing" nursing homes experienced a 31 percent improvement in the CMS quality measure while the remaining "non-high implementing" nursing homes experienced a 12 percent decline in the same measure.
- The intervention's effectiveness was limited to the subset of nursing homes that researchers found had most fully involved staff at all levels in adopting the new workflow and technology.

According to the methodological literature and experts that we consulted, certain information can provide the basis for an assessment of the consistency in an intervention's effects on quality of care and costs in different organizations. Specifically, this information concerns the number of different organizations where the intervention has been implemented, the degree of diversity exhibited by those organizations, and the consistency in observed changes in quality of care and costs across those organizations. However, such information would not be available for assessing the consistency of results across diverse organizations if an intervention has been implemented in only a few different organizations, or in multiple organizations that are generally quite similar. That is also the case if studies only analyze and report changes in quality of care and costs attributed to an intervention in the aggregate, rather than separately for the different organizations that implemented it.

On the other hand, for interventions that have been implemented in multiple, diverse organizations, and their results analyzed separately at the different organizations, it is possible for policymakers to compare the results of the intervention across those organizations to examine the consistency of the intervention's effect. To the extent that those interventions consistently produce positive effects on quality of care and costs among diverse organizations, that provides evidence of their capacity for widespread replication. For other interventions, if data on the changes in quality of care and costs across the different organizations indicate a lack of consistency in outcomes, that provides evidence of a more restricted capacity for replication.

Leadership Support and Other Factors Reported As Important for Both Implementation and Replication of Interventions

Respondents to our questionnaire reported, generally by large margins, that leadership support as well as other factors, such as organizational culture and staff resources, significantly facilitated implementation. However, respondents were more divided when asked about the reported effect that health IT had on implementation, and most respondents reported that financial incentives were not a factor in the implementation of their intervention. A majority of respondents reported that each of these factors, with the exception of financial incentives, was expected to be either very or somewhat important if one were to attempt to replicate their intervention as widely as possible.

Most Respondents Reported Leadership Support and Other Factors Significantly Facilitated Implementation of Interventions

Taking account of factors that prior research has shown tend to facilitate or impede the implementation and replication of interventions may enhance efforts by policymakers and others to promote the adoption of interventions across varied organizational contexts. In examining the relative impact of seven factors identified in our literature review, we found that respondents to our questionnaire reported, generally by large margins, that five of the seven factors significantly facilitated implementation of their intervention. Health IT and financial incentives were the exceptions. Leadership support was the factor that the largest number of respondents reported as having significantly facilitated implementation of their intervention (see table 4). When asked to describe how leadership support facilitated implementation, respondents frequently explained that a leader who visibly

prioritized and endorsed the intervention, allocated necessary resources, and championed the development and implementation of the intervention and drove necessary organizational or behavioral changes facilitated the implementation of the intervention. Respondents also explained that having champions, specifically clinicians, was a key factor in encouraging cooperation and participation in the intervention by staff, especially fellow clinicians. The prominent role attributed to leadership in implementing the many different types of interventions in our sample suggests that policymakers will have greater success in implementing and replicating interventions to the extent that they can take steps to ensure that strong leadership is in place before interventions are initiated.

Table 4: Reported Effect of Identified Factors on Implementation of Selected Health Care Interventions

Factors identified from relevant literature	Significantly facilitated	Somewhat facilitated	Somewhat impeded	Significantly impeded	Not a factor	Don't know/ No basis to judge	Total
	(number of respondents)						
Leadership support	92	22	2	3	2	4	125
Availability of tools and activities to standardize care	72	33	3	0	11	6	125
Staff resources	68	29	14	2	10	2	125
Organizational culture	60	35	13	6	8	3	125
Financial resources	50	27	15	5	23	5	125
Health IT	29	30	14	10	36	5	124
Financial incentives	20	18	1	2	68	15	124

Source: GAO analysis of responses to GAO questionnaire.

Note: Table excludes missing responses.

Respondents typically reported that a combination of additional factors along with leadership support significantly facilitated implementation of their intervention. The 92 respondents who reported leadership support as having significantly facilitated implementation, reported, on average, another three factors as having significantly facilitated implementation. Of the 86 respondents who reported at least one factor in addition to leadership support as significantly facilitating implementation, more than half reported staff resources (60), organizational culture (55), and the availability of other tools (50), respectively, as having significantly facilitated implementation. Nearly half (42) reported

that financial resources, in addition to leadership, significantly facilitated implementation. Just six respondents reported leadership support and no other factor as having significantly facilitated implementation.

In contrast to the five factors that a clear majority of respondents reported having facilitated implementation of their intervention, respondents were more divided on how health IT affected implementation, as shown in table 4. Health IT had the highest number of respondents, compared to the other factors, that reported the factor impeded implementation of their intervention.³³ Further, a substantial group of respondents reported that health IT was not a factor. On the other hand, nearly half of respondents reported that health IT either significantly or somewhat facilitated implementation of their intervention. Respondents frequently explained that health IT facilitated implementation of their intervention by enhancing the exchange of information and communication across providers or organizations, facilitating the collection of data or the evaluation of the intervention and improving the efficiency and productivity of staff. Of those who reported that health IT impeded implementation, respondents commonly cited the limited functional capacity of existing systems or the lack of interoperability across settings as impediments to successful implementation.³⁴ Other respondents explained that the general lack of health IT altogether acted as a barrier that impeded implementation. Variation in the role of health IT across different types of interventions does not appear to explain the mixed assessment of this factor; as respondents for each of the intervention types included in our sample—with two exceptions—were similarly divided on how health IT affected implementation. However, proportionately more respondents for care coordination or transitions of care interventions as well as care-process-

³³ Respondents who reported a factor as having impeded implementation of their intervention—either somewhat or significantly—commonly explained that the lack or insufficiency of the factor acted as a barrier that impeded implementation.

³⁴ Health IT is interoperable when systems are able to exchange data accurately, effectively, securely, and consistently with different IT systems, software applications, and networks in such a way that the clinical or operational purposes and meaning of the data are preserved and unaltered.

improvement interventions reported health IT as having facilitated implementation compared to respondents for other types of interventions.³⁵ This result suggests that as policymakers consider different health care interventions, implementation of some of their options will depend more heavily than others on having appropriately configured health IT in place.

Financial incentives were most often reported as not a factor. Slightly more than half of our 127 respondents reported financial incentives—as distinct from the related, but broader, financial resources factor—as having not been a factor in implementation of their intervention. The exception was for the two types of interventions for which financial incentives were an integral component—provider payment restructuring and insurance redesign—where respondents most often reported financial incentives as having significantly facilitated implementation. When asked to explain how financial incentives facilitated or impeded implementation most respondents simply provided a description of the incentives they used to implement the intervention, such as payments to providers or patients to participate in the intervention. However, a few respondents explained that the expected cost savings generated from the intervention was an indirect incentive to implement the intervention while other respondents explained that incentives within existing payment systems, or the lack thereof, affected implementation. While the implementation of many interventions included in our sample may not have been affected by financial incentives, current means of paying for health care, such as fee-for-service payment structures, may have hindered the successful implementation of other interventions.

³⁵Care coordination or transitions of care interventions set out to facilitate patient transfers from one setting to another. Some focus on coordination of patient care delivered by multiple providers.

Most Respondents Expected Leadership Support and Other Factors Would Be Very Important for Widespread Replication

Much as they had reported regarding the implementation of their intervention, nearly all respondents consistently expected that leadership support would be very important if one were to attempt to replicate their intervention as widely as possible (see table 5). Leadership support was reported nearly unanimously by respondents as being very important for widespread replication of their intervention, paralleling respondents' relatively consistent assessment of the effect of leadership on implementation. In addition, a clear majority of respondents expected that each of the other factors—except for financial incentives—would be either very or somewhat important for replication.

Table 5: Expected Degree of Importance of Identified Factors for Widespread Replication of Selected Health Care Interventions

Factors identified from relevant literature	Very important	Somewhat important	Not important	Don't know/No basis to judge	Total
		(number of respondents)			
Leadership support	116	7	0	0	123
Staff resources	79	41	2	1	123
Organizational culture	79	39	1	2	121
Availability of tools and activities to standardize care	71	41	9	2	123
Financial resources	66	44	11	2	123
Health IT	48	48	23	4	123
Financial incentives	30	23	53	15	121

Source: GAO analysis responses to GAO questionnaire.

Note: Table excludes missing responses.

In contrast to the highly divided views health IT evoked from respondents regarding its role in the implementation of their interventions, it was reported by a substantial majority of respondents as either very (48) or somewhat (48) important for widespread replication. This could be an indication that, if health-IT-related impediments experienced when implementing the intervention, such as the lack of interoperability across settings, were ameliorated, health IT could be important to the successful replication of some interventions. Similar to views expressed about the implementation of care coordination or transitions of care interventions, respondents for these types of interventions commonly reported that health IT would be very important for

widespread replication more so than respondents for other types of interventions.

Financial incentives was the factor that drew the most mixed assessment from respondents with regards to its expected importance for the widespread replication of interventions. Nearly half of respondents indicated that financial incentives were not important for widespread replication, which is similar to the view of most respondents regarding the role of such incentives in the implementation of their interventions. Another substantial group of respondents (30) indicated that financial incentives would be very important for replication. When respondents were asked to explain why factors would be important for widespread replication, respondents discussed financial factors more frequently than any other factor. Respondents' explanations about these financial factors often concerned a misalignment of financial incentives within existing payment systems that limited the attractiveness of replicating interventions that seek to enhance value. For example, some respondents noted that it would be difficult to replicate interventions that involved providing additional services, such as care coordination, under existing payment systems that typically do not compensate providers for those services.

Concluding Observations

Our work suggests that progress in achieving greater value in health care in the U.S. will depend, in part, on the availability of information regarding the effect of different interventions on quality of care and costs and on how policymakers and others assess and use that information. Such information can guide the choices of policymakers among multiple interventions vying for support, but those decisions will have a sounder basis if the information meets certain criteria regarding its content and strength of evidence. With respect to content, information on the magnitude of an intervention's effect on both quality of care and costs is needed to determine if an intervention has enhanced value. In the case of the responses to our questionnaire on 127 diverse interventions, we found that this basic level of information was reported as available about half the time. With respect to the strength of evidence, the most critical indication comes from the types of study designs used to produce that information. There are a range of rigorous study designs which can provide credible support for the attribution of

observed changes in quality of care and costs to a particular intervention. Our review of studies associated with the 127 interventions examined by our questionnaire found that while a number of studies employed rigorous study designs, a substantially larger number employed weaker designs that could not isolate the effect of an intervention from other factors. To the extent that policymakers find and use information on health care interventions that provides sufficient credible evidence on the effects of those interventions on both quality of care and costs, they will be better equipped to determine which interventions produce greater value in health care. Our work also suggests that successful efforts to encourage the widespread adoption of value-enhancing interventions will need to take into account a complex mix of factors, including leadership support, organizational culture, and staff resources, that facilitate the implementation of health care interventions across a wide range of organizational contexts.

Agency Comments

We requested comments from the Department of Health and Human Services, but none were provided.

We are sending copies of this report to the Secretary of Health and Human Services and other interested parties. In addition, the report will be available at no charge on the GAO Web site at <http://www.gao.gov>.

If you or your staff have any questions regarding this report, please contact me at (202) 512-7114 or cosgrovej@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix V.



James Cosgrove
Director, Health Care

Appendix I: Scope and Methodology

To examine the availability of information on the effect of selected health care interventions on quality of care and costs as well as factors that can facilitate the implementation and replication of these interventions, we studied a diverse set of specific interventions that seek to enhance the value of health care through making changes in the way care is delivered. Specifically, these interventions make changes in who delivers health care services, how care is organized, or where care is delivered for a specified population. To identify interventions for our study, we drew upon six distinct sources to select a broad and diverse, though not exhaustive, set of interventions that have been implemented in one or more locations in the U.S. or abroad. These sources allowed us to identify a wide range of value-enhancing strategies implemented in different of health care settings, such as hospitals, integrated delivery systems, and physician practices, over more than a 10-year time span, including interventions that have not been described in academic or professional literature. We identified 828 interventions of potential relevance to our study through the following six sources:

- A review of relevant literature on health care interventions that make changes in who delivers care services, how it is delivered, or where it is delivered. We conducted several searches using online databases, including Medline and ProQuest Health, to identify articles on interventions that were published from 1999 to 2009.
- A review of interventions contained in the Agency for Healthcare Research and Quality's (AHRQ) Health Care Innovations Exchange (HCIE) as of August 20, 2009. The HCIE is a Web site that acts as a repository for information on quality improvement interventions and other innovative strategies to improve health care submitted by their implementers. Many of the interventions contained in the HCIE make changes in the way health care is delivered and include information on cost as well as quality.¹

¹See www.innovations.ahrq.gov for more information on the HCIE, last accessed on February 23, 2011.

- A review of the relevant articles contained in the *Tufts Medical Center's Cost-Effectiveness Analysis (CEA) Registry* that were published from 1999 to 2009. The CEA Registry is a comprehensive database of health care cost-utility analyses that examine the health benefits and costs of strategies to improve health care. The CEA Registry contains articles from 45 peer-reviewed publications.²
- Interviews with experts on health care interventions associated with organizations such as state governments, integrated delivery systems, employer groups, and other countries.
- Information on interventions that we identified from press reports, select journal articles published after 2009, and presentations at conferences.
- Information on interventions submitted by their innovators or evaluators to either the Senate Budget Committee or GAO.

To select interventions for inclusion in our study, we reviewed source documents for each of the potentially relevant interventions that we identified through our six sources. We selected 239 interventions that met the following seven criteria:

- The intervention made a discrete change in who delivers health care services, how care is organized, or where care is delivered.
- The intervention targeted a population or problem that was relevant to the U.S. health care system.³
- The intervention may have included health information technology (health IT) as one of its components of change, but health IT was not the intervention's only component of change.

²See www.cearegistry.org for more information on the CEA Registry, last accessed on June 2, 2011.

³For example, a treatment program to improve care for children and adults with malaria in sub-Saharan Africa was excluded from our set of interventions. Although this intervention was instituted to improve and standardize the delivery of health care services, malaria is not a disease commonly identified in the U.S. population.

- The primary goal of the intervention was not focused on increasing access to care.
- The intervention activities must fall within the health care system.⁴
- The source document or documents for the intervention either contained information or indicated that information is available on the effect of the intervention on quality of care and its effect on costs. Moreover, the source documents indicated that the intervention enhanced the value of health care by meeting one of the following three conditions: (1) increases quality of care and reduces costs; (2) maintains quality of care and reduces costs; or (3) increases quality of care and maintains costs.
- The intervention was implemented in at least one health care setting. Interventions that were studied by examining their potential costs and benefits based on simulated outcomes rather than analyzing data from their actual implementation were excluded.

To collect information on the 239 health care interventions that we selected for our study, we developed a Web-based questionnaire that contained 22 open- and closed-ended questions on interventions, their effect on quality of care and costs, and factors that may affect their implementation and replication. We sent our questionnaire to 235 individuals who participated in developing, implementing, or evaluating each intervention.⁵ We identified these individuals through the source documents that we used to select interventions for our study. We received usable responses—responses that contained relevant information on the effect of the intervention on quality of care, the effect of the intervention on costs, or key factors that may affect implementation—for 127

⁴For example, a community-based prevention program to limit alcohol consumption at licensed premises was excluded from our set of interventions. This intervention made changes in the community rather than within the health care system to reduce alcohol-related health problems.

⁵Contacts for four of our selected interventions declined to participate in our questionnaire.

interventions. We developed protocols for cleaning and analyzing data that we received from questionnaire respondents.⁶ These protocols included: identifying usable responses; reviewing source documents to clarify responses; and, if necessary, contacting respondents directly to obtain additional information on their intervention.

To determine the availability of information on the effect of selected health care interventions on quality of care, we analyzed data that we collected from respondents through our questionnaire. We asked respondents to describe up to five key measures used to assess the effect of their intervention on quality of care and the magnitude—a percentage change or other quantitative assessment—of change observed in each measure described relative to a control group that did not experience the intervention or a baseline assessment made prior to implementing the intervention. We conducted a content analysis on questionnaire responses to determine the number of respondents who described one or more key measures used to assess the effect of their intervention on quality of care and the number of respondents who reported improvements in those measures attributable to their intervention.

As part of our analysis on the availability of information on the effect of selected health care interventions on quality of care, we examined the types of quality measures respondents reported. We conducted a content analysis on questionnaire responses to determine what aspect of care quality—such as patient mortality, hospital readmissions, or patient satisfaction with care—each measure examined. We categorized each measure that respondents described by type based on the aspect of care quality it examined; for example, we categorized a measure that assessed the effect of an intervention on patient mortality as an outcome measure. We categorized quality measures into types that are largely based on the measure domains laid out by AHRQ in its

⁶Our analysis of factors that facilitate or impede implementation and replication were based on 125 responses. Two out of 127 interventions for which we received usable responses did not answer any questions on key factors for implementation.

National Quality Measure Clearing House.⁷ We did not include all measure domains laid out by AHRQ in our analysis, because some domains, such as access to care, fell outside of the scope of our engagement. Moreover, measures that did not clearly specify which aspect of care quality was assessed were categorized as unspecified measures. We analyzed this information to determine the types of quality measures used to assess the effect of each intervention on quality of care.

To determine the availability of information on the effect of selected health care interventions on costs, we analyzed data that we collected from respondents through our questionnaire. We asked respondents to report the type of cost savings, such as total dollars saved or dollars saved per patient, calculated to assess the effect of their intervention on costs and the specific amount saved for type of cost savings calculated. We also asked respondents if their reported savings accounted for costs associated with implementing the intervention and what information was used to calculate those savings. We determined the number of respondents who reported calculating each type of savings and a specific amount saved for those savings. Furthermore, we analyzed responses by finding the number of respondents who reported accounting for costs associated with implementing the intervention—net cost savings—and the type of information used to calculate those savings. Additionally, we used this information along with information we obtained through our analysis of quality measures to determine (1) the number of respondents who reported a magnitude of improvement in quality measures and a specific amount saved attributable to their intervention and (2) the number of respondents who also reported net cost savings rather than gross cost savings.

To identify key criteria that can be used to assess the strength of available evidence on the capacity of interventions to enhance the value of health care we interviewed methodological experts and

⁷The National Quality Measure Clearinghouse identifies seven measure domains to classify the focus of a quality measure. AHRQ's measure domains: access of care, outcome of care, patient experience of care, population health, process of care, structure of care, and use of services. See <http://qualitymeasures.ahrq.gov/> to access AHRQ's National Quality Measure Clearinghouse. Last accessed on June 22, 2011.

conducted a literature review to identify relevant systems for assessing the strength of evidence. We reviewed methodological literature published by entities that have well-established systems for evaluating health care interventions, including the Cochrane Collaboration; AHRQ's Effective Health Care Program, which includes the Evidence-based Practice Centers; and in the United Kingdom, the National Institute for Health and Clinical Excellence and the Centre for Reviews and Dissemination. We focused on those entities with systems for evaluating organizational interventions that change the structure or delivery of health care. This led us to pay particular attention to the guidance developed by Cochrane's Effective Practice and Organisation of Care (EPOC) Group, a collaborative review group that specializes in conducting systematic reviews of organizational interventions.⁸

Our review of this methodological literature and guidance together with our expert interviews led us to develop a set of questions to help decision makers and policy analysts who support them to critically examine the strengths and limitations of evidence about health care interventions that seek to enhance value. These questions target three broad areas: (1) assessing the true effect of the intervention on quality of care and costs, (2) assessing the scope of study results, and (3) assessing an intervention's capacity for replication. We submitted our initial draft questions to several different experts in assessing the comparative effectiveness of health care interventions and received their feedback on the content and clarity of those questions. Based on that feedback, we made revisions, resulting in the criteria described in our report and the set of questions listed in appendix IV.

As part of our efforts to identify key criteria for assessing the strengths and limitations of available evidence on the capacity of interventions to enhance the value of health care, we examined the

⁸Other assessment approaches that we examined, but found less relevant for organizational interventions, were the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system and the U.S. Preventive Services Task Force. See G.H. Guyatt et al, "GRADE: an emerging consensus on rating quality of evidence and strength of recommendations," *BMJ*, 336 (April 2008), 924-926; Agency for Healthcare Research and Quality, *U.S. Preventive Services Task Force Procedure Manual*, AHRQ Pub. No. 08-05118-EF, July 2008.

choice of study design used by evaluators to study the interventions for which we received usable responses to our questionnaire. To determine the type of study design used to assess the effect of interventions, we reviewed source documents and questionnaire responses. (See app. III for more information on study designs.) Some interventions reported results from multiple studies. In these cases, we identified each type of study design used to assess the intervention. We used this information to find the number of interventions that were assessed using more rigorous study designs such as randomized controlled trials and the number of interventions that were assessed using less rigorous study designs such as pre/post or cross sectional studies.

Our approach is designed to assist decision makers and policy analysts in assessing the strengths and limitations of evidence provided to them about the effects of health care interventions on quality of care and costs. Our approach does not involve the performance of systematic reviews that could synthesize information about those effects from multiple studies. Nor does it attempt to describe a process for producing a numerical or qualitative rating of the methodological strength of a study along one or more specified dimensions. Rather, our approach emphasizes the questions that decision makers and policy analysts should ask and leaves open the format and content of the answers to those questions.

To examine factors that can facilitate the implementation and replication of health care interventions that seek to enhance value, we analyzed data collected from respondents through our questionnaire. We reviewed key literature sources and interviewed experts to identify seven factors that may affect implementation including leadership support, organizational culture, and

resources.⁹ Respondents were asked to indicate, from the list of close-ended categorical options, to what degree each of the seven factors facilitated or impeded implementation and to provide an open-ended explanation of how the factors facilitated or impeded implementation. We asked respondents who were familiar with the replication of their intervention to explain if and how the factors differed from site to site. Respondents were also asked to indicate the expected degree of importance that each factor could have in attempting to replicate the intervention as widely as possible and to explain why these factors were expected to be important. In addition to the factors identified through our literature review, we asked respondents to identify and describe up to three additional factors that facilitated or impeded implementation of their intervention or that would be important for wide-scale replication. All close-ended responses were analyzed by assessing the frequency distribution of responses for each factor. We conducted a content analysis on open-ended responses to identify common explanations of how these factors affected implementation and why these factors would be important for widespread replication of the intervention.

⁹J. Ovretveit. *Does improving quality save money? A review of evidence of which improvements to quality reduce costs to health services providers*, (London: The Health Foundation, September 2009); C. VanDeusen Lukas et al, "Transformational change in health care systems: An organizational model," *Health Care Management Review*, vol. 32, no. 4 (2007); C. Homer and R. Baron, "How to Scale Up Primary Care Transformation: What We Know and What We Need to Know?" *Journal of General Internal Medicine*, vol. 25, no. 6 (2010); M. Wang et al, "Redesigning Health Systems for Quality: Lessons from Emerging Practices," *Joint Commission Journal on Quality and Patient Safety*, vol. 32, no. 11 (2006); T. Greenhalgh, G. Robert, F. MacFarlane, P. Bate, and O. Kyriakidou, "Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations," *The Milbank Quarterly*, vol. 82, no. 4 (2004); E. Ferlie and S. Shortell, "Improving the Quality of Health Care in the United Kingdom and the United States: A framework for Change," *The Milbank Quarterly*, vol. 79, no. 2 (2001); S. Shortell et al, "Assessing the Impact of Continuous Quality Improvement/Total Quality Management: Concept versus Implementation," *HSR: Health Services Research*, vol. 30, no. 2 (1995); T. Bodenheimer, "The American Health Care System: The Movement for Improved Quality in Health Care," *New England Journal of Medicine*, vol. 340, no. 6 (1999); and E. Bradley et al, "The Roles of Senior Management in Quality Improvement Efforts: What are the Key Components?," *Journal of Healthcare Management*, vol. 48, no. 1 (2003).

As part of our analysis of factors that may affect implementation and replication, we examined differences in questionnaire responses by the intervention type. To determine the types of interventions for which we received usable questionnaire responses, we reviewed source documents and questionnaire responses for each intervention and assigned them to one of eight categories (see app. II for more information about intervention type). To categorize interventions by type we assessed key intervention characteristics, including the population targeted for behavior change and levers or activities used to change the way health care services are delivered. For example, a hospital surgical team that implemented a checklist was categorized as a *patient safety improvement intervention*. Some interventions exhibited key characteristics of more than one type of intervention. For example, a primary care practice that implemented a nurse case manager to facilitate care transitions and employ disease management strategies exhibits key characteristics of both *care coordination or transition of care programs* and *chronic condition management* interventions. Interventions that exhibited key characteristics of more than one type of intervention were categorized in all appropriate types. To determine if the effect or expected degree of importance of the factors differed by the type of intervention, we assessed the frequency distribution of responses for each factor across intervention type.

Although our efforts to identify relevant interventions for our study were extensive, we could not ensure that every intervention meeting our selection criteria had been identified. Therefore the results from our questionnaire are limited in scope to the 127 interventions for which we received usable responses, and cannot be generalized to all value-enhancing health care interventions.

Appendix II: Types of Health Care Interventions That Seek to Improve the Value of Health Care

Intervention type (number in study)	Description	Examples
Provider payment restructuring (3)	<ul style="list-style-type: none"> Interventions that seek to alter provider behavior by systematically changing the basis for provider payments. 	<ul style="list-style-type: none"> Providing a single payment, or bundled payment, for all health care services that are delivered for a defined episode of care or a specified period of time. Providing physician group practices performance payments if the practice meets or exceeds performance targets.
Insurance redesign (6)	<ul style="list-style-type: none"> Interventions that seek to alter patient behavior by restructuring health insurance plan provisions or related health care benefits. 	<ul style="list-style-type: none"> Insurers offer enrollees a tiered network of providers. Enrollees who choose a provider in the higher cost tier pay higher premiums or cost sharing than enrollees who choose a provider in a lower cost tier. Enrollees are charged a lower or no copay for specific drugs that are part of a recommended medical regimen for a medical condition.
Chronic condition management (38)	<ul style="list-style-type: none"> Interventions that seek to improve care for patients with chronic conditions. Can be implemented in either inpatient or outpatient settings. Can focus on patient or clinician activities, or both. 	<ul style="list-style-type: none"> A nurse-social worker team is introduced into a primary care practice to provide education, help patients improve self management skills, and develop care plans with patients. A multidisciplinary team holds classes for children with severe asthma and their parents to address physical needs and group, individual and family therapy for psychological needs.
Patient safety improvement (26)	<ul style="list-style-type: none"> Interventions that seek to prevent or reduce adverse events caused by medical care. Adverse events include improper prescriptions or administration of medications, health-care associated infections, and pressure sores. 	<ul style="list-style-type: none"> A surgical team implements a check list that enhances team communication and situational awareness among clinicians to prevent wrong-site surgeries. A program of patient risk assessments, specialist consultations, and new equipment is designed to minimize pressure sores.
Care coordination or transition of care programs (24)	<ul style="list-style-type: none"> Interventions that facilitate patient transfers from one setting to another. Some focus on coordination of patient care provided by multiple providers. 	<ul style="list-style-type: none"> An advanced practice nurse and a trained elder peer provide support to older adults who are discharged home after a heart attack or undergoing bypass surgery to encourage compliance with medications and lifestyle changes. A team of nurses and social workers work with patients with multiple chronic conditions to coordinate care from multiple providers and to provide ongoing monitoring and referrals.
Continuous system improvement (3)	<ul style="list-style-type: none"> Interventions that seek to change health care organization as a whole through ongoing and iterative reassessment of health care practices. Such interventions seek to both reduce inefficiency or waste and improve patient outcomes. 	<ul style="list-style-type: none"> A hospital created teams trained in “lean” principles, based on Toyota’s manufacturing approach, to identify where changes in routine procedures could reduce waste and increase efficiency.

Appendix II: Types of Health Care Interventions That Seek to Improve the Value of Health Care

Intervention type (number in study)	Description	Examples
Prevention (4)	<ul style="list-style-type: none"> The primary goal is to improve health by forestalling the development of illness in the first place. 	<ul style="list-style-type: none"> Programs to promote wellness activities and health screenings or to prevent falls. These interventions do not include programs to prevent adverse events.
Care process improvement (31)	<ul style="list-style-type: none"> Interventions that seek to ensure that clinical staff adhere to specified treatment protocols or other forms of standardized practices. These interventions seek to modify care processes by changing where care is delivered, how care is organized or structured, or who delivers care. 	<ul style="list-style-type: none"> Multi-site intensive care unit telemedicine program. A team of clinicians use a four-step mobility protocol to regularly assess the functional and clinical status of intensive care unit patients with respiratory failure.

Source: GAO.

Appendix III: What Makes Some Study Designs More Rigorous Than Others

The methodological literature on assessing the effect of interventions places a major emphasis on study design for identifying those studies that have the capacity to assess an intervention's effect on an outcome.¹ The key strength of rigorous study designs is that they can take account of other factors that could affect the outcome of interest, and thereby isolate the effect of the intervention itself.

Randomized controlled trials (RCTs) are widely considered to be among the most rigorous types of study designs because their basic structure inherently minimizes the potential impact of confounding factors on their results. RCTs accomplish this by randomly allocating study participants to groups that either receive the intervention—generally referred to as intervention or treatment groups—or do not receive the intervention—the control groups. The consequence of random allocation is that the only systematic difference between study participants in the two groups is exposure to the intervention.² Thus, the effect of all other factors is the same on the two groups and therefore neutralized in making comparisons between the intervention and control groups.

A second design type, known as the controlled before and after study, can be used in situations where the random allocation of study participants between intervention and control groups required for an RCT is not feasible. Controlled before and after studies use data collected from separate treatment and control groups, both before and after the intervention's implementation, to help to separate the effect of the intervention from that of other factors at work over that time period. In this design type, the control group is generally chosen in a way that is likely to produce a group that is broadly similar to the treatment group prior to the implementation of the intervention. However, methodologists generally recommend an explicit analysis to compare the intervention and control groups

¹See for example WR Shadish, TD Cook, and DT Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, (Boston: Houghton Mifflin Co., 2002).

²Systematic differences involve patterns or trends. Where there are no systematic differences, all that is left is random variation.

used in controlled before and after studies in order to demonstrate that they were in fact similar before the intervention took place.

A third design type, an interrupted time series study, is not based on a comparison of intervention and control groups.³ Instead, it tracks an outcome of interest over time with measurements taken at many different time points both before and after the intervention. The multiple data points from before the implementation of the intervention enable analysts to take account of the impact of other factors on the outcome and thereby isolate the intervention's effect on that outcome. The interrupted time series design works best when there are data from a substantial number of different time points, both before and after implementation of the intervention.⁴

Other types of study designs cannot isolate the effect of an intervention from that of other factors because they provide no separate information on what would have happened without the intervention. For example, in a simple pre/post study all one has is a measurement of the outcome before implementation of the intervention and a measurement of the outcome after the intervention. The observed difference reflects all the factors (including the intervention) affecting the outcome over that time period. Because confounding factors could potentially affect the outcome in either the same or the opposite direction as the intervention, the actual effect of the intervention itself could be either greater or smaller than the simple pre/post difference. Even the direction of the intervention's effect, to increase or decrease the outcome, could be the opposite of the overall change from pre to post. That is why the results of a pre/post study generally cannot be relied on to provide even an approximation of what the likely effect of a health care intervention is on quality of care and costs.⁵

³However, it is possible to add a control group to an interrupted times series design to enhance its capacity to isolate the effect of an intervention.

⁴There are other rigorous study design types that can be used in certain situations. One example is the regression discontinuity design that allocates study participants to either an intervention or control group based on whether they have a score above or below a specified cut-off point on an assignment variable.

⁵An exception would be the unusual situation where it was established that no other factors besides the intervention could have affected the outcome over that time period.

Appendix IV: Key Questions for Assessing Evidence from Studies of Interventions That Seek to Enhance Value

The following three tables provide a set of questions that are intended to help policymakers and others find the information needed to assess the strengths and limitations of evidence drawn from studies of health care interventions that seek to enhance value relating to their impact on quality of care and costs.¹ The three tables focus on the three broad dimensions described in the body of this report: (1) the credibility of evidence that attributes changes in quality of care and costs to the intervention, (2) the applicability of study results for broader populations of interest, and (3) the intervention's capacity for widespread replication.

Each table lists a series of questions that highlight key information for assessing the evidence produced by relevant studies along with guidance on how to look for that information in published reports. Answers to most of these questions may be found in relevant sections of those reports;² if not, one can ask the investigators who conducted the studies. While this set of questions is selective and does not cover every potential methodological issue, the information it calls for should provide policymakers a basis for making an informed assessment of the overall credibility and scope of the available evidence regarding the apparent impact of these interventions on quality of care and costs, as well as the demonstrated capacity of those interventions for widespread replication.

¹Health care interventions enhance value when they lower or hold costs constant while sustaining or increasing quality of care.

²These could include sections that describe study methods and results (e.g., characteristics of study participants), as well as sections describing the implications of study results including their limitations.

**Appendix IV: Key Questions for Assessing
Evidence from Studies of Interventions That
Seek to Enhance Value**

Table 6: To Assess the Credibility of Attributing Observed Changes in Quality of Care and Costs to the Intervention

Key question	Guidance on finding and interpreting information needed
1. Did the study use a rigorous design type?	<ul style="list-style-type: none"> • Determine from study descriptions of methodology if the study used a rigorous study design type, such as: <ul style="list-style-type: none"> • randomized controlled trials • interrupted times series • controlled before and after study • Or if it used a non-rigorous design type, such as: <ul style="list-style-type: none"> • pre/post study • post-only study
2. Were the intervention and control groups similar?	<ul style="list-style-type: none"> • The two groups were similar if a study: <ul style="list-style-type: none"> • used random allocation to create the treatment and control groups, or • analyzed how well the treatment and control groups matched on major characteristics prior to implementation of the intervention and found no major differences.
3. Were changes in quality assessed using appropriate measures?	<ul style="list-style-type: none"> • Determine if the study cites other research documenting that the validity and reliability of the quality of care measures used in the study had previously been tested or to demonstrate that the measures had frequently been used in related research.
4. Were changes in costs assessed in a way that took account of implementation costs?	<ul style="list-style-type: none"> • In calculating the cost effect of the intervention, studies should subtract all the costs of implementing the intervention, including both start-up costs and ongoing costs. • Return on investment is another way of presenting information on the cost effects of an intervention.^a
5. Were data collected consistently?	<ul style="list-style-type: none"> • Examine study methodology for descriptions of procedures to ensure that the same data were collected from different groups in the study—such as intervention and control groups—at the same time and in the same way. • Determine if the study considered potential vulnerabilities to consistent data collection and took steps to mitigate them, such as blinded assessment of outcomes.^b
6. Were the data collected sufficiently complete?	<ul style="list-style-type: none"> • Studies should report the proportion of participants that dropped out of the study and therefore provided either no or partial data. • Concerns about data completeness are raised if <ul style="list-style-type: none"> • the overall proportion of study participants who dropped out of the study exceeds 20 percent. • the proportion of study participants who dropped out of the treatment and control groups is not broadly similar.

Source: GAO.

^aReturn on investment (ROI) for an intervention represents the ratio of the change in overall costs of care attributed to the intervention divided by the cost of implementing the intervention. A positive ROI would have a value greater than one. An ROI between 0 and 1 means that savings attributed to the intervention were less than the cost of implementing the intervention. An ROI of less than 0 indicates that the intervention led to increased costs.

^bBlinded assessment of outcomes means that those who collect data on quality outcomes do not know which study participants were assigned to the treatment and control groups.

**Appendix IV: Key Questions for Assessing
Evidence from Studies of Interventions That
Seek to Enhance Value**

Table 7: To Assess the Applicability of Study Results for Broader Populations of Interest

Key Question	Guidance on finding and interpreting information needed
1. From what larger group or groups are study subjects, also known as the study population, chosen?	<ul style="list-style-type: none"> • Examine description of study methodology to identify the group or groups from which study participants were selected (also known as a universe). • Identify the key characteristics that defined the universe and thereby determined who or what was eligible for inclusion in the study, such as type of provider.
2. What method or mechanisms are used to select study participants?	<ul style="list-style-type: none"> • If study descriptions indicate that the participants were selected randomly from a larger defined group, then this random selection makes the study population representative of the universe from which it was selected.
3. Do analyses show that the study population and any broader populations of interest are broadly similar on key characteristics? ^a	<ul style="list-style-type: none"> • Discussion or comment sections may analyze key similarities and differences between the study population and broader populations of interest. • To the extent that analyses of the study population compared with that of a population of interest establish that the two populations are broadly similar in terms of their key characteristics, those analyses support the representativeness of the study population for that broader population. • To the extent that analyses of the salient characteristics of the study population compared with that of a population of interest find that the two populations differ on one or more key characteristics, those analyses diminish support for the representativeness of the study population for that broader population. • If a study population differs from a broader population of interest on a given characteristic, it may still be representative of that broader population if persuasive evidence is presented to show that the characteristic on which the two differ does not affect performance of the intervention.

Source: GAO.

^aKey characteristics are those that are likely to affect the performance of the intervention. Determining which characteristics are key may involve a mix of judgment, reference to past research, and analysis of study data.

**Appendix IV: Key Questions for Assessing
Evidence from Studies of Interventions That
Seek to Enhance Value**

Table 8: To Assess an Intervention’s Capacity for Widespread Replication

Key Question	Guidance on finding and interpreting information needed
1. In how many different organizations has the intervention been tested?	<ul style="list-style-type: none"> • Determine if study reports indicate the number of different locations or sites at which the intervention was tested. • Different physical locations or sites represent different organizations to the extent that the success of an intervention in one site is independent of its success in another.
2. How diverse were the organizations in which the intervention has been tested?	<ul style="list-style-type: none"> • Determine if study reports describe the characteristics of the locations or sites at which the intervention was tested. • The more varied the organizations in which the intervention is tested, the stronger the test of its ability to be replicated successfully in different organizational contexts. • The more detailed the description of the sites, the more complete is the information on the extent of diversity among them.
3. How uniform has implementation of the intervention been in different organizations?	<ul style="list-style-type: none"> • Determine from study descriptions the extent to which the intervention accommodates variation in how it is implemented in different organizations. • If variation across organizations is substantial, identify, if possible, a subset of organizations that implemented essentially similar versions of the intervention.
4. How consistent are the intervention’s results across the different organizations where the intervention was implemented?	<ul style="list-style-type: none"> • Examine study results for analyses of the extent of variation across different organizations included in the study that implemented similar versions of the intervention.^a • An intervention could have consistent results on quality measures but not costs, or vice versa.

Source: GAO.

^aIf not provided in published studies, disaggregated results for individual organizations can be requested from the researchers who conducted assessments of the intervention.

Appendix V: GAO Contact and Staff Acknowledgments

GAO Contact

James Cosgrove (202)-512-7114 or cosgrovej@gao.gov

Acknowledgments

In addition to the individual named above, Jessica Farb, Assistant Director; Kristin Ekelund; Krister Friday; Katie Mack; and Eric Peterson made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday afternoon, GAO posts on its Web site newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to www.gao.gov and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's Web site, <http://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Ralph Dawn, Managing Director, dawnr@gao.gov, (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

