Chapter 4

Issues in Designing an Evaluation of PPS

Contents

FIGURE

Chapter 4

Issues in Designing an Evaluation of PPS

INTRODUCTION

This chapter examines issues that arise in developing a strategy for evaluating the impact of Medicare's prospective payment system (PPS). These issues are of two kinds:

- fundamental choices that must be made with respect to the kinds of questions that are asked about PPS; and
- tradeoffs that are necessary in the selection of specific research designs for answering the questions.

In this report, the term "evaluation" as applied to PPS refers to any effort to associate changes in characteristics of the health care system with the implementation of PPS or its components. The usefulness of an evaluation in guiding policy and program changes varies directly with the quality of evidence on which such associations are based and the confidence with which causal inferences can be made.

FUNDAMENTAL ISSUES UNDERLYING THE DEVELOPMENT OF A PPS EVALUATION PLAN

Important choices are necessary before critical PPS evaluation questions can be developed. These choices involve specifying the kinds of information that are important in guiding policy and program changes. OTA made such choices in developing the critical evaluation questions summarized in chapter 1. Considerations underlying these decisions are discussed below.

Specifying the Perspective of Evaluation

One of the first questions to be addressed in developing a PPS evaluation plan is whose perspective is important in the evaluation. The impacts of PPS on Medicare expenditures was, of course, critical to the passage of the law establishing PPS, and these effects need to be assessed. But such a narrow program perspective is inadequate in evaluating a program with the wide range of effects of PPS.

PPS will have varying effects on Medicare beneficiaries, other cohorts of the population, providers of health care, suppliers of medical products, employees, educators, and researchers. An evaluation could proceed from the perspective of one or a combination of these groups. A comprehensive evaluation would be one that would balance the effects on these different groups and take the interests of society as a whole into account. This report recognizes the tradeoffs among affected groups and lays out evaluation questions regarding the distribution of such effects across members of society.

Specifying the Standard of Comparison

A second question involves the standard against which PPS effects will be judged. PPS contains three central elements: 1) a system of expenditure control carried forward from the Tax Equity and Fiscal Responsibility Act (TEFRA) (Public Law 97-248); 2) a restructuring of financial incentives from cost-based reimbursement to per-case payment; and 3) the use of diagnosis-related groups (DRGs) to classify patients for the purposes of payment.

Some of the effects of PPS might occur in any system that controlled revenues going to hospitals. If one wanted to compare the results of Medicare's PPS with the results of alternative systems of expenditure control, one could analyze evidence in the four States—Maryland, Massachusetts, New Jersey, and New York—currently hold-

38-747 0 85 - 3 : QL 3

ing waivers from Medicare's PPS. [1] As noted in chapter 3, the ratesetting systems in these States are different from PPS in major and minor aspects of program design. The State systems are generally all-payer systems, the unit of payment in some is not the admission, and the method of arriving at the payment rates differs in each State. The State systems are required by their waiver contracts, however, to hold expenditures to levels that are no higher than those that would have occurred under PPS. Consequently, comparative analyses of experience under PPS and under the State waivers can be highly revealing of the specific gains and losses from a system like PPS compared to other kinds of expenditure control.

For immediate policy, understanding the impacts of particular elements of PPS, such as the special treatment of teaching hospitals or the use of DRGs as the patient classification system, may be even more important than understanding the overall effects of PPS relative to cost-based hospital reimbursement. Yet it is nearly impossible to identify the specific aspects of PPS that are responsible for any observed changes in the health care system. Indeed, as the remainder of this report shows, attributing any changes in the health care system to PPS as a whole will be difficult enough. Thus, it is probably infeasible to evaluate the impact of specific components of PPS on the behavior and outcomes of the health care system. In general, then, this report deals with approaches to evaluating the effects of PPS as a whole relative to Medicare's former cost-based hospital reimbursement system.

One component of PPS that is particularly critical to the incentives in the system is the use of DRGs as the system of classifying patients for payment purposes. Other approaches to patient classification have been or are under development, and DRGs themselves are likely to be refined as time goes by. (The DRG and alternative patient classification systems are described in app. H.)

Under Medicare's DRG-based PPS, the financial desirability of any given patient is established in part by his or her assignment to a particular DRG. Patient classification systems other than the DRG system would group patients in different ways, changing the amounts paid for, and therefore the relative profitability of, some patients. The patient classification system can also influence the financial incentives involving the use of existing medical technologies and the introduction of new ones. Patient classification systems such as DRGs, which assign patients at least partly on the basis of whether a specific technology is used, can encourage or discourage the introduction of new technologies whose use would change the patient's category (see ch. 8 for examples). Classification systems that assign patients to categories on the basis of clinical condition alone, not resource use, create very different incentives for new technologies, The focus of this study is on the incentives inherent in Medicare's DRG-based PPS relative to cost-based reimbursement, but comparative studies of the effect of alternative approaches to patient classification on hospital behavior regarding admissions strategies and technology use and adoption would be useful as well.

Defining the Objectives of Evaluation

A third major issue is the relative importance of various desirable features of an evaluation of PPS. One has to address the tradeoffs among the following, partially competing, objectives:

- to act as an early-warning system for serious, unintended consequences of PPS;
- to obtain a balanced view of PPS effects, including both positive and negative impacts;
- to quantify any observed effect with precision;
- to attribute any observed effect to PPS with confidence;
- to afford the research effort; and
- to choose feasible evaluation approaches.

The affordability and feasibility of using particular impact measures and research methods are determined largely by the kinds and quality of data sources available, the cost of obtaining the data, and the administrative or ethical barriers to their use, Feasibility is also limited by the lack of comprehensive and balanced measures of quality of care (see ch. 6) and access to care (see ch. 7). In the meantime, the measures that do exist,

I Together, these four States account for 533, or 9 percent, of the non-Federal community hospitals in the United States (13).

if chosen carefully, may give an acceptably accurate picture of how the health care system is changing in this regard.

The feasibility of attributing observed effects to PPS is limited by several factors. One problem is that because PPS has been implemented universally among non-Federal community hospitals (except in the four States with waivers), the opportunities for comparison are limited. Another problem is that PPS is not the only change underway in the U.S. health care system; simultaneous influences, which can often be accounted for only by the passage of time, confound attempts to directly attribute many changes in the health care system to PPS.

Despite such difficulties, it is still possible to conduct pre/post-PPS analyses that offer strong suggestive evidence about the impacts of PPS or its components. Success hinges on careful a priori analysis of the likely magnitude and direction of influence of other factors so that the effects of PPS may be reasonably well inferred.

Determining the Evaluation Time Schedule

An evaluation plan must take into account the fact that the effects of PPS will unfold only over time, as the health care system gradually adapts to the new payment environment. Some changes may occur early and continue throughout the lifetime of PPS; some may occur early but disappear as PPS goes on; others may not surface until much later. Reductions in personnel staffing levels, for example, appear to take place almost immediately as hospitals have sought quick responses to the incentives of PPS (379), but these changes may not be long-lived. Or, the incentive to move patients out of the hospital early to reduce length of stay may not be acted on in the short-run if facilities to care for these patients are in short supply. **As** time goes by, however, if the health care system responds with an increase in the supply of long-term care facilities, shifts in the settings of care may be more dramatic.

One of the most fundamental changes encouraged by PPS is also likely to take a number of years to occur. Physicians' attitudes may gradually become more positive toward the appropriateness of taking cost into consideration in clinical decisionmaking (275). Increasing interest in issues of cost-effectiveness of medical practices will lead to more research into these questions and ultimately more information available to physicians. Yet, these developments are likely to be quite gradual, showing their influence on patterns of medical care utilization only after years.

A strategy for monitoring and evaluating PPS should take account of the timing of effects as well as the ultimate impacts. Certain observable changes in the health care system may be able to serve as valid early warning indicators of important long-run effects of PPS; the challenge is to choose them correctly when little evidence of such validity is available.

ALTERNATIVE EVALUATION RESEARCH DESIGNS

A question that transcends all areas of impact is how to design an evaluation that will provide sufficiently valid answers about the impacts of PPS at reasonable cost and in a timely manner, A range of evaluation designs can be considered, each with its own strengths and weaknesses.

In selecting an evaluation design, one must consider the potential validity of the findings against the cost (or, alternatively, the affordability) and the feasibility of the approach. The validity of a study is defined here as the extent to which explanations other than the program under study can be ruled out as responsible for the observed effect (internal validity); and the extent to which the findings can be generalized beyond the study sample (external validity) (54). Often, the level of validity obtainable with a particular research design varies directly with cost and inversely with feasibility.

Figure 4-1 summarizes the performance of alternative evaluation designs on the dimensions of feasibility, affordability, and validity. Controlled

Figure 4-1.–Comparison of the Feasibility, Affordability, and Validity of Alternative Designs for Evaluating the Impact of PPS on the Health Care System

Feasibility/ affordability	Validity		
	Low	Medium	High
Low		Studies using comparison groups	Controlled random experiments
Medium		Pre/post-program comparisons	
High	Opinion Surveys, anecdotes	Case studies	

SOURCE Off Ice of Technology Assessment, 1985

experiments typically have high validity (particularly internal validity) because they are carried out prospectively and generally involve the random assignment of subjects to an experimental program or to a control group (or program). Since both groups are exposed to whatever simultaneous influences occur, differences in study outcomes can reasonably be ascribed to the experimental program. Unfortunately, program evaluations can rarely take place in such an environment. In the case of PPS, the program has been implemented universally, with waivered States generally unrepresentative of the rest of the country.

Quasi-experimental research designs move back from the strict requirements of controlled experiments to the use of ***comparison groups*** whose representativeness has not been established or to ***pre/post-program comparisons.*** Only if the analyst has a high level of confidence that the comparison groups are likely to be representative or that observed effects are unlikely to be due to simultaneous influences can these designs offer much validity.

More informal approaches, such as detailed ***case studies,*** are systematic efforts to identify behavior or outcomes that can be linked in very specific circumstances to the program under study. Case studies of decisionmaking in hospitals, for example, can uncover behavior that is in direct response to PPS. Case studies of the development and diffusion of medical technologies provide an excellent means of identifying aspects of the payment system that affect technological change (see ch. 8 for more detail). These approaches to evaluation can be enlightening, but they pose threats to validity that need careful attention. Most important, bias in the selection of subjects for case studies, a phenomenon that is difficult to guard against, can call attention to some effects of PPS and ignore others that are equally important.

As chapter *3* of this report illustrates, much of the information currently available on the impacts of PPS consists of ***anecdotes arid opinions*** (sometimes systematically collected through surveys). Often, observed changes in a measure of effect may be ascribed to PPS on the basis of opinion. For example, changes in employment patterns in hospitals since PPS have been documented. Whether or to what extent such changes are due to the implementation of PPS is unknown, however. A survey might ask hospital administrators for their opinions regarding the importance of PPS relative to other factors in bringing about these changes.

Evidence consisting of anecdotes and opinions has the lowest validity for obvious reasons—the opinions may be biased and the anecdotes rare outliers—but it is not necessarily wrong. In fact, such informal sources of data can be most useful as early warning systems that raise hypotheses about the impacts of PPS. Reliance on these sources without further analysis, however, exposes policymaking to high risks.

ASPECTS OF PPS AFFECTING RESEARCH DESIGN

The most appropriate research design depends on the particular characteristics of PPS that affect the feasibility, cost, and validity of the alternatives laid out above. Three such characteristics are as follows:

- PPS is a moving target;
- PPS is being implemented in an environment of multiple and major simultaneous influences; and
- the availability of data for some kinds of designs is limited.

Each of these aspects is discussed below.

PPS as a Moving Target

A problem common to all evaluations of major programs is that the character of the program itself changes over its lifetime. Evaluations of the effects of a program in its early years may be irrelevant by the time the evaluations are finished, In the case of PPS, major changes are inevitable, The 3-year phase-in period institutionalizes a policy of change. Not only are hospitals finding their revenues increasingly subject to PPS, but the DRG prices themselves are moving from a regional to national basis. Moreover, important components of hospital costs, namely capital and direct medical education expenses, have been excluded from PPS, but there is reason to believe that these exclusions will not persist in the next 5 years. Finally, the overall generosity of PPS, which is largely determined by the annual rate of increase in the average DRG price, may have more to do with impacts on the health care system than any other aspect of the program and is subject to variation over time as cost-containment pressures grow or recede.

These realities argue for a continuous system of monitoring PPS effects that focuses on sensitive and readily available indicators of system performance such as changes in patterns of expenditures, utilization, and organization of care. But information on these indicators needs to be buttressed by research linking them to the important, ultimate impacts on health care benefits and costs.

Simultaneous Influences in the Health Care System

Were PPS the only change underway in the health care system, it would be possible to compare outcomes before and after its imposition in order to infer its impacts. But the health care system has been undergoing rapid change in the past 5 years and continues to be dynamic. Among the most important simultaneous influences are the following:

- ***The ratio of physicians to population has been increasing.*** The supply of physicians has been rising dramatically in the past decade as a consequence both of Federal policy on medical education and the immigration of foreign-trained physicians. In 1975, there were 179 physicians per 100,000 people in the United States. By 1981, this number had grown to 207 (346). The trend toward higher physician-to-population ratios ***is*** expected to continue, with an expected ratio of 264 per 100,000 people by the year 2000 (346).
- ***Competition in the health care system has been increasing.*** Numerous factors have contributed to an increase in the amount of competition for patients. First, the supply of physicians and innovative health care facilities has increased in the past 5 years and continues to increase. New alternative sites of health care delivery, such as freestanding ambulatory surgical and emergency centers, have been formed throughout the country. For example, between 1979 and 1982, the number of freestanding emergency centers grew from 44 to almost 500 nationwide (292). Hospitals have attempted to compete with freestanding facilities by upgrading hospital-based emergency rooms and providing their own freestanding facilities (292).

  Second, increasing pressure on employers to contain the costs of their health benefits has led to changes in health insurance plans, which encourage competition among providers on the basis of price. Increased beneficiary cost-sharing requirements, for exam-

pie, make patients more price sensitive in selecting their settings of care. The development of preferred provider organizations, which contract with insurers to provide services at a reduced rate, is the latest manifestation of increasing competition (307).

Third, the size of the uninsured population tends to fluctuate with the business cycle but has been on the rise in recent years (283). With more patients lacking insurance, they are likely to become more sensitive to hospitals' prices and hospitalizations are delayed.

- ***Various aspects of Federal health policies have undergone changes concurrent with PPS.*** Other aspects of Federal health policy have been altered during the time immediately prior to or during the implementation of PPS. For example, the Medicaid program, which is administered by States under general guidelines and financial subsidies from the Federal Government, has given increasing flexibility to the States to define eligibility, scope of covered services, and levels of payments to providers. In fiscal year 1982, after the passage of the Omnibus Budget Reconciliation Act of 1981 (Public Law 97-35), the number of Medicaid recipients per capita and Medicaid payments per recipient declined (111). These declines occurred despite a severe economic recession that would be expected to raise welfare rolls and Medicaid expenses.

  The Veterans Administration (VA) is also currently undergoing changes in the way funds for patient care will be allocated among facilities. The VA has begun to implement a new budget allocation system that will tie budgets more closely to standardized work units. For inpatient admissions, the DRG definitions are being used. Facility budgets will come to be more dependent on case mix than in the past, when high occupancy was rewarded regardless of turnover. The new budget systems provide incentives to VA facility managers for cost containment and more selective admission and treatment criteria (110). Consequently, it may be difficult to separate the effects of PPS on the utilization of VA services from the effects of the VA's own administrative changes.

  The Medicare program itself has undergone substantial changes in policy concurrent with PPS. For example, the Deficit Reduction Act of 1984 (Public Law 98-369) mandated a fee schedule for ambulatory laboratory procedures, including those performed by hospital laboratories, that may reduce the incentives for hospitals to provide laboratory services to hospital outpatients and physician office practices. The imposition of a freeze on physician fees and changes in Medicare assignment policies also alter the environment in which PPS operates, although the directions and extent of their effect are unknown.

- ***State policies have changed concurrently with*** PPS. 'As mentioned above, individual-States have acted to contain the costs of their Medicaid programs in a variety of ways (307). State certificate-of-need programs, intended to constrain the supply of hospital and nursing home beds and expensive capital equipment, have waxed and waned with State politics (308). In a few instances, States have enacted laws to make the environment more favorable to competition for health care. California, for example, passed a law in 1982 which permitted the formation of preferred provider organizations, which enhances price competition in the health care market (79). In 1984, an estimated 15 to 20 percent of the State's 25 million people were served by preferred provider organizations (370).

Taken together, these simultaneous influences substantially reduce the validity of pre/post comparisons of measures of effect. They also suggest that differences among States and regions of the country in the health care environment will jeopardize the validity of State-by-State comparisons of health system variables. Only those effects with the strongest hypothesized direct link to PPS can be analyzed in such a way, and even then imperfectly.

For example, all experts agree that PPS should shorten the average length of stay (ALOS) in short-term hospitals. ALOS has been falling gradually for both Medicare and non-Medicare patients over the past 5 years (see ch. 3). One can extrapolate from this previous trend to predict

changes in ALOS that would have occurred in the absence of PPS and then compare these predictions with actual ALOS since the beginning of PPS. ALOS may also be compared among hospitals with varying shares of Medicare patients. But neither of these approaches is entirely valid. We cannot know with certainty whether the predicted trend in ALOS is an accurate representation of what would have occurred in the absence of PPS or whether the ALOS might have shifted one way or another on its own. Yet the demonstration of a significant shift in ALOS from previous trends concurrent with or shortly after the implementation of PPS remains strongly suggestive that PPS is having the expected kinds of effects on a critical measure of hospital utilization. Thus, imperfect as the evidence of PPS on ALOS is, it provides an approximate estimate of PPS effects that needs to be linked to the more important questions of PPS impacts on cost, quality, and access.

Data Availability

The choice of research design is inextricably related to the kinds of data that are available and the costs of making necessary data available. The use of data routinely collected by the Health Care Financing Administration to administer the Medicare and Medicaid programs or of data available from organizations with ongoing surveys clearly offers cost advantages over special surveys or other primary data collection methods. Because these systems have not been developed or maintained with an eye to their usefulness as tools for program evaluation, however, they omit important data elements, and some items are so unreliable that analysis cannot proceed.

Moreover, the content and reliability of the data change over time, complicating pre/post-PPS comparisons. For example, prior to 1982, the assignment of diagnostic and procedural codes to Medicare hospital claims data was sloppy, because payment was not based on these items and the data entry procedures were inadequate (213). This information is expected to improve markedly for post-PPS years, but other data items that are not important for payment may deteriorate in quality. [2] Comparisons of billing claims in the post-PPS era with those in the pre-PPS period may be complicated by this problem.

Conversely, pre/post-PPS comparisons of the content of medical care delivered in hospitals would require detailed review and abstracting of medical records, an approach to data collection that is reasonably reliable but very costly (see ch. 6). And observed changes in impact measures could still not be ascribed to PPS with complete confidence.

[2] For example, PPS puts a new premium on rapid submission of claims following a patient discharge. Hospitals may not have an incentive to code discharge status accurately.

CONCLUSIONS

Medicare's PPS is a complex program instituted in an even more complex health care environment. As a radical new approach to hospital payment, it needs to be evaluated for its impacts on health care costs, quality of care, access to care, technological change in medicine, and clinical research. Yet it is important to be realistic about what can be expected from such evaluations.

A variety of research designs are potentially applicable to the evaluation, each with its own strengths and weaknesses. The tradeoff of validity with cost and feasibility is critical to optimal selection of research designs from among the alternatives.

Because the availability of data figures so directly in the choice of impact measures and research designs, the temptation is great to study only those questions that are easy to study because of data availability. The danger of this situation is that it may result in an unbalanced view of the impacts of PPS. Consequently, it is important to identify at the outset the critical evaluation questions that need to be addressed. Serious consideration needs to be given to ways of ad-

dressing each of these critical questions, using methods and data that reflect the tradeoff between validity, feasibility, and affordability. Part Two of this report examines each of the five PPS impact areas identified in chapter l—health care, expenditures and costs, quality of care, access to care, technological change, and clinical research —and identifies the critical evaluation questions in each. It also examines the data available to support analyses of these questions and suggests specific studies that appear to be worth their costs.