United States General Accounting Office

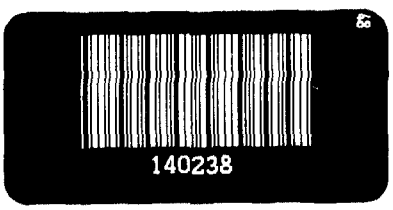
GAO

Report to the Chairman, Subcommittee on Health, Committee on Ways and Means, House of Representatives

December 1989

MEDICARE

Improvements Needed in the Identification of Inappropriate Hospital Care



GAO/PEMD-90-7

GAO

United States
General Accounting Office
Washington, D.C. 20548

Program Evaluation and
Methodology Division

B-234152

December 20, 1989

The Honorable Pete Stark
Chairman, Subcommittee on Health
Committee on Ways and Means
House of Representatives

Dear Mr. Chairman:

In response to a request from your subcommittee, we are submitting this report on approaches to assessing the appropriateness of hospital care. In particular, the report (1) examines available information on the extent of inappropriate care in the Medicare program, (2) describes the current approaches used by Medicare and the private sector to review the appropriateness of hospital care, and (3) suggests approaches that might reduce the level of inappropriate care in the Medicare program. This report contains a recommendation that the Health Care Financing Administration (HCFA) implement expanded prospective review based on systematic studies of the cost-effectiveness of various options.

We are sending copies of this report to the Secretary of the Department of Health and Human Services and interested congressional committees and members, and will make copies available to others upon request. If you have any questions or would like additional information, please call me at (202) 275-1854 or Dr. Lois-ellin Datta, Director of Program Evaluation in Human Services Areas, at (202) 275-1370. Other major contributors to this report are listed in appendix IV.

Sincerely yours,

Eleanor Chelimsky
Assistant Comptroller General

Executive Summary

Purpose

The Medicare program is the most important avenue through which over 31 million elderly people obtain hospital care in this country. Payments for inpatient hospital services in fiscal year 1988 totaled $51.9 billion. The Health Care Financing Administration (HCFA) is responsible for assuring that Medicare beneficiaries receive appropriate medical care of high quality. GAO has previously examined HCFA's efforts to assure quality of care in the Medicare program. This report focuses on assessing the appropriateness of hospital care, which GAO defines in terms of the need for specific services and the proper setting and intensity of those services. Issues of quality of care, such as how well the medical care was provided and whether some needed treatment was omitted, are beyond the scope of this report. Requested by the Subcommittee on Health of the House Committee on Ways and Means, this report has three objectives:

- to examine critically the available information on the extent of inappropriate care in the Medicare program,
- to describe what is currently being done in both the Medicare program and the private sector to review the appropriateness of medical care, and
- to suggest approaches that might be effective in reducing the level of inappropriate care in the Medicare program and issues that would have to be addressed if HCFA were to adopt those alternatives.

Background

Both public and private insurers have developed programs to examine the appropriateness of the individual medical services covered by their health care benefits plans. These programs are generically referred to as utilization review. In the Medicare program, the 54 Utilization and Quality Control Peer Review Organizations (PROs) are responsible, under contract to HCFA, for judging the appropriateness of hospital care provided to Medicare beneficiaries, as well as other aspects of the care, including its quality and the accuracy of hospital classification decisions affecting reimbursement levels. Utilization review activities in the private sector are conducted by insurance companies, large employers, managed care organizations such as health maintenance organizations and preferred provider organizations, and a variety of companies that conduct utilization review on a contractual basis.

The decision-making process in utilization review typically involves an examination of medical records or other patient information by nurse reviewers using explicit (written) review protocols. Cases identified by a nurse as possibly inappropriate admissions are referred to a physician.

The physician, applying his or her own knowledge, expertise, and experience, makes the final determination of appropriateness based on the medical record and any additional information obtained from the attending physician.

Results in Brief

GAO found that PRO reviews of hospital care have typically identified a lower rate of inappropriate care than have reviews conducted by SuperPRO (an independent HCFA contractor) or by researchers. GAO also found a number of differences between PRO and alternative reviews that might explain the discrepancy. These include the criteria used to screen cases, the cases selected for review, and the lack of incentives for PROs to aggressively question the appropriateness of care. All tend to decrease the rate of inappropriate hospital care uncovered by the PRO reviews.

The primary difference between private sector and PRO utilization review activities is timing. Private sector programs operate prospectively, focusing on the identification, prior to admission, of patients who do not require hospitalization and their diversion to more appropriate health care settings. Except for a small number of prospectively reviewed surgical procedures, most PRO reviews are carried out retrospectively—that is, after the patient has been discharged from the hospital. There are also important differences in the type of cases targeted for review, the nature of determinations that the proposed care is inappropriate, and the structure of benefit packages.

Principal Findings

Available Evidence Suggests More Inappropriate Care Than Reported by PROs

Between 1986 and 1988, 2.1 percent of Medicare hospital admissions reviewed retrospectively by PROs were denied for being inappropriate. SuperPRO reviewed a random sample of Medicare cases previously reviewed by PROs and questioned the appropriateness of 12 percent of the admissions. In research studies of inappropriate hospital admissions, between 7 and 19 percent of admissions were judged to be inappropriate.

For specific procedures prospectively reviewed by PROs, the denial rate varied from 0 to 11.5 percent, depending on the procedure. Higher denial rates were generally associated with surgical procedures that can

be performed on an outpatient basis. Estimates of inappropriate procedures from research studies varied from 14 percent to 32 percent. Independent reviews, conducted for GAO by experienced medical reviewers, of 213 coronary angiographies performed on Medicare beneficiaries in fiscal year 1987 found that 20 percent were inappropriate, a result very similar to estimates published previously by researchers.

PRO and Private Sector Utilization Review Activities Differ

GAO found that the primary difference between private sector and PRO utilization review activities is that the private sector emphasizes prospective reviews, which occur before the attending physician has expended time and resources on the proposed care. This permits the attending and review physicians to discuss potential alternatives and agree on a course of treatment. In PRO retrospective reviews, the interaction between the PRO reviewers and the attending physicians involves questioning the appropriateness of past actions. PROs must balance the need for positive working relationships with their local medical communities with the denial of payment for care determined to be inappropriate. Thus, it may be more difficult for PRO reviewers than for the private sector to aggressively question the appropriateness of care.

Three related aspects of the private sector approach also seem important. First, private prospective reviews target potentially inappropriate hospital admissions, typically all elective admissions. Most PRO retrospective reviews do not target cases with high rates of inappropriate admissions. Second, the private sector appropriateness decision is "advisory"—that is, the physician and patient are advised that the proposed care cannot be certified, but that a payment decision will not be made until a bill is submitted. This allows the individual and his or her physician to make the final decision. PROs have no comparable option. Finally, many private health benefit plans provide a penalty for not complying with the requirement to obtain preadmission certification, or an incentive for complying with the utilization review advice, or both. The current structure of Medicare benefits would not permit this use of penalties and incentives.

The limited information on the cost-effectiveness of private sector prospective review clearly indicates that reductions in hospital use and expenditures more than offset the costs of conducting the reviews. However, because the Medicare population is older and probably in poorer health than the general population, the cost-effectiveness of particular options for Medicare prospective reviews would need to be considered.

Some Changes to the PRO Program Would Require New Legislative Authority

HCFA's current authority would allow the expanded use of prospective review and better targeting of the reviews. However, PROs do not currently have the ability to make prospective decisions that are "advisory," nor is the program structured in such a way as to provide penalties (other than outright payment denial) or incentives to encourage compliance with prospective review decisions. These changes would require new legislative authority.

Changing the emphasis in Medicare's appropriateness review from retrospective to prospective would require some reevaluation of the structure of retrospective reviews. For example, instead of reviewing individual admissions, the reviews might focus on identifying providers with unusual practice patterns for more intensive review and on verifying information provided during prospective review.

Recommendation

GAO recommends that the Secretary of Health and Human Services require the Administrator of HCFA to expand prospective utilization review in the Medicare program. This should be preceded by a systematic examination of the cost-effectiveness of various options, including those described in this report. If HCFA lacks legislative authority to implement particular options, such as "advisory" prospective determinations and the use of incentives or penalties for complying with prospective review decisions, they should seek that authority from the Congress. It is also important to evaluate the current retrospective review requirements in light of expanded prospective reviews, to avoid unnecessary duplication and improve the efficiency and effectiveness of both types of review.

Agency Comments

The Department generally agreed with GAO's recommendation to expand Medicare prospective review. They contend that they have already expanded prospective review in the third PRO Scope of Work and will examine the review results and change the program as needed. GAO is not convinced that the changes in HCFA's third PRO Scope of Work will demonstrate the potential cost-effectiveness of prospective review because few of the procedures reviewed are ones that should normally be treated in an outpatient setting. In addition, no attention is given to medical cases for which hospital admission typically is not warranted. Such cases would be most likely to yield cost savings in prospective reviews. The Department should also conduct systematic studies of the cost-effectiveness of other options—such as the use of "advisory" review decisions and contracting with firms specializing in appropriateness review—and then adopt cost-effective options.

Contents

Tables

Abbreviations

AEP	Appropriateness Evaluation Protocol
DRG	Diagnosis-related group
GAO	General Accounting Office
HCFA	Health Care Financing Administration
ISD-A	Intensity of Service, Severity of Illness, Discharge Screens, and Appropriateness review system
PPS	Prospective Payment System
PRO	Utilization and Quality Control Peer Review Organization
PROMPTS	Peer Review Organization Monitoring Protocol and Tracking System
PSRO	Professional Standards Review Organization
UR	Utilization review

Introduction

Both public and private insurers have developed programs to examine the appropriateness of the medical services provided to individuals covered by their health care benefits plans. These programs are generically referred to as utilization review (UR). The impetus driving the development of these programs has been primarily cost—reflecting a concern that the availability of health care coverage might lead to the provision of unnecessary, and costly, health services. However, the actual review of appropriateness focuses on whether the care is necessary by evaluating the risks and benefits of particular services for particular types of patients.

One of the primary goals of UR is a reduction in the amount of inappropriate care provided to patients. As a result of the reduction in the amount of inappropriate care, health care payors expect to see a related reduction in their expenditures. The reduction also saves patients from the physical and financial risks associated with inappropriate care. Finally, UR programs seek to "educate" physicians and others providing inappropriate care about their inappropriate practice patterns.

Hospital UR in the Medicare Program

Medicare pays much of the health care costs for the 31 million eligible people aged 65 and older and for some of the disabled. Hospital care is covered under Part A, Hospital Insurance, and financed primarily by Social Security payroll taxes. Payments for inpatient hospital services in fiscal year 1988 totaled $51.9 billion.

Although some public sector UR programs existed prior to 1965, a major expansion of UR activities occurred as a result of the 1965 Medicare and Medicaid legislation, which included mandatory provisions for hospitals to establish UR committees. Amendments to the Social Security Act in 1972 established Professional Standards Review Organizations (PSROs) to monitor the medical care provided to Medicare beneficiaries, thereby creating a review entity independent of the hospitals. Beginning with the passage of the Tax Equity and Fiscal Responsibility Act of 1982, responsibility for review of the appropriateness, as well as the quality, of care was passed to the newly created Utilization and Quality Control Peer Review Organizations (PROs).

Within the Department of Health and Human Services, the Health Care Financing Administration (HCFA) is responsible for the overall administration of Medicare, including establishing the regulations and policies under which the PRO program operates. PROs perform their review activities for HCFA under contractual arrangements, which are either renewed

or competitively rebid on three-year cycles. As of May 1989, all PROs are in the third cycle of contracts; 44 organizations have contracts covering the 54 states and territories.

Two other groups of organizations also conduct related UR activities under contract to HCFA. An entity known as "SuperPRO" assesses the adequacy of PRO determinations concerning the appropriateness and quality of inpatient care and also the validity of diagnosis-related group (DRG) assignments.[1] In addition, the fiscal intermediaries—either commercial insurance companies or Blue Cross plans—pay claims and make adjustments in payments as directed by either HCFA or PROs. Their activities include checking bills for completeness, assuring that the patient is entitled to the services, and adjusting payments to providers based on the results of PRO and other reviews.

Hospital UR in the Private Sector

UR activities in the private sector are conducted by insurance companies, large employers, managed care organizations such as health maintenance organizations and preferred provider organizations, and a variety of companies that conduct UR on a contractual basis. In recent years, as a response to escalating health expenditures, there has been a dramatic increase in the private sector UR activities. A 1988 listing of 95 firms offering contractual UR services indicated that 74 percent had begun providing UR services in the preceding five years. A September 1988 survey of 100 benefits managers of large corporations found that approximately half the firms began UR activities less than three years previously.

Despite the growth of private sector UR programs, relatively little systematic information is available about their program components or effectiveness. It is in this context that the Subcommittee on Health of the House Committee on Ways and Means requested that we examine currently available private sector approaches for reviewing the appropriateness of medical care and determine whether they are suitable for use in the PRO program.

[1] About 90 percent of the payments for services provided to hospital patients are made under the Prospective Payment System (PPS). Under PPS, hospitals receive payments based on predetermined rates for 475 different groupings of diagnoses and procedures. The groupings are referred to as DRGs.

Objectives, Scope, and Methodology

Objectives

In response to the Subcommittee's request, this report has three objectives:

- to examine critically the available information on the extent of inappropriate care in the Medicare program,
- to describe what is currently being done in both the Medicare program and the private sector to review the appropriateness of medical care, and
- to suggest approaches that might be effective in reducing the level of inappropriate care in the Medicare program and issues that would have to be addressed if HCFA were to adopt those alternatives.

Scope

This report focuses on approaches for reviewing the appropriateness of hospital care. In keeping with the general health services research literature, we will use the term "appropriateness" to refer to two broad categories of issues:[2]

- Are, or were, the particular services needed?
- If so, in what setting should the services be, or have been, provided, and for how long?[3]

The first question logically precedes the second and asks whether the proposed diagnostic and/or therapeutic services are, or were, needed by the patient, given his or her unique circumstances. For example, is it appropriate to do coronary artery bypass graft surgery on a 50-year-old man with single vessel disease and moderate angina? Does a patient with benign hypertrophy of the prostate gland need immediate surgery? Does a patient with cataracts who cannot walk unaided and lives in a nursing home need intraocular lenses?

[2]A. Donabedian, Explorations in Quality Assessment and Monitoring, Vol. 1, The Definition of Quality and Approaches to Its Assessment (Ann Arbor, Mich.: Health Administration Press, 1980); S.M.C. Payne, "Identifying and Managing Inappropriate Hospital Utilization: A Policy Synthesis," Health Services Research, 22:5 (December 1987), pp. 709-69.

[3]The term "services," in this report, refers to the entire range of medical and surgical services that could potentially be delivered in a hospital. However, utilization review of individual hospital services typically focuses on invasive diagnostic and surgical "procedures." In later sections of this report, we often use the more specific term "procedure" in place of the more general "services."

The second question involves two judgments. First, does, or did, the patient's condition (that is, the severity of the illness) or the service(s) provided (that is, their intensity and required professional skills and backup support) warrant hospitalization? Second, if hospitalization is, or was, warranted, how long a stay is, or was, needed? For example, does a diabetic require hospitalization to be restabilized on insulin and, if so, for how long? How frail does a person have to be before cataract surgery should be done as an inpatient, rather than an outpatient, procedure? How long should a person stay in the hospital, on average, after having a heart attack?

This focus on the appropriateness of a decision to provide particular health services to a particular patient excludes consideration of a number of other issues. First of all, the outcome of care is not a consideration in our examination of appropriateness. One might conclude, for example, that a particular operation is medically appropriate, even if some patients die during the operation, because their chances of dying are greater if they do not have the operation. Second, our report focuses on appropriateness in terms of what was or was not done for the patient, rather than on how well it was done. The issue of how well something was done is a quality-of-care issue beyond the scope of this report. Third, an important aspect of UR excluded from this report is the validation of the information regarding the patient's diagnoses and the services provided. Both PROs and the private sector engage in these kinds of reviews. Finally, while we critically examine different estimates of the extent of inappropriate care in the Medicare program for comparative purposes, we do not attempt to reach conclusions about the actual level of inappropriate care.

HCFA's Definition of Appropriateness

The Medicare program is prohibited by statute from paying for medical services that "are not reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member."[4] The terms "reasonable and necessary" are not defined in current HCFA regulations. However, in January 1989, HCFA published a proposed rule intended, at least in part, to define these terms and to "assure that federal funds are expended only for medical services that are appropriate [emphasis added] to meet an individual's medical needs" (54FR4303). An appreciation of how HCFA operationalizes these terms is important for understanding the role of PROs in reviewing the appropriateness of care.

[4]Quoted from section 1862(a)(1)(A) of the Social Security Act.

Under the proposed regulations, a medical service would be considered reasonable and necessary (and thus covered) if it were safe and effective, neither experimental nor investigational, appropriate, and cost-effective. HCFA notes that not all of these criteria will be applicable to every coverage question and that safety and effectiveness are the key factors. These criteria apply to the effectiveness of individual medical services in general and not to their appropriateness for individual patients.

In addition to these general criteria, under the proposed regulations, Medicare reviewers, including PROs, may consider whether the service is

- furnished in accordance with accepted standards of medical practice,
- medically necessary in the particular cases and for the duration and frequency of its use or application, and
- furnished in a setting appropriate to the patient's medical needs and condition.

The three foregoing issues are most relevant to our study in that they specify the criteria that may be considered in reviewing the medical care provided to specific patients, rather than general rules for judging the overall effectiveness of a particular medical service. In particular, the last two criteria correspond quite closely to the two appropriateness issues presented in the previous section of this report.

PROs are contractually required by the Third Scope of Work to review retrospectively the hospital care provided to selected cases in six basic categories, only some of which pertain to appropriateness. These review categories are quality review, discharge review, admission review, invasive procedure review, coverage review, and DRG validation. An examination of the specifications for these reviews makes it clear that only the admission review and the invasive procedure review relate primarily to issues of appropriateness. The admission review is intended to detect inappropriate admissions. The invasive procedure review requires PROs to review, for appropriateness, all surgical procedures and any other procedure that would affect DRG assignment. PROs are also required to review, on a preadmission/preprocedure basis, the appropriateness of ten surgical procedures.[5] These appropriateness reviews are examined in greater detail in subsequent chapters.

[5]The retrospective invasive procedure review required under the third PRO Scope of Work should not be confused with the preadmission reviews of appropriateness that the PROs are also required to perform.

Methodology

The information presented in this report comes primarily from the published literature and interviews with individuals involved in the review of medical care. In addition, we analyzed available data on reviews of the appropriateness of care provided to Medicare beneficiaries. We also selected a sample of medical records for Medicare patients who had had coronary angiographies and had them independently reviewed to determine the appropriateness of the procedure.

Our work was performed in accordance with generally accepted government auditing standards. We requested and received formal written comments from the Department of Health and Human Services. Their comments, and our responses, are summarized in chapter 4 and reproduced in appendix III.

Literature Review

The primary method that we used to identify approaches for examining appropriateness was a search of the published literature. We conducted a bibliographic search using key words such as "medical necessity" and "inappropriate care." We also extended the literature search to include citations regarding preadmission certification programs, as well as prior approval or second opinion programs for elective surgery. We obtained the studies identified in that search, as well as others cited in their bibliographies. We also inquired about relevant literature during our interviews.

Interviews

To learn more about specific identified approaches, we contacted over 60 individuals in organizations involved in medical care appropriateness issues. These included academic researchers, the developers of particular protocols identified in our literature review, and a number of health care provider and insurance organizations. We talked with individuals involved in private sector review of appropriateness, including representatives of UR firms, insurance companies, and managed care entities such as health maintenance organizations and preferred provider organizations. Finally, we interviewed HCFA program officials responsible for UR, as well as officials of three PROs.

Secondary Data Analyses

Our secondary data analyses were aimed at generating additional information on levels of inappropriate care and at examining potential explanations for any differences. First, we used information on PRO admission review activities that HCFA periodically tabulates from monthly reports submitted by individual PROs. PROs vary considerably in regard to when they submit these monthly reports to HCFA. In addition, HCFA sometimes requires revisions to the reports before they are accepted. As a result, the precise time period covered by these data is very difficult to specify

because it varies by individual PRO.[6] However, it covers roughly the period from June 1986 to February 1989. Data from this source are referenced as HCFA data and provide rates of admission denial for reviews of the appropriateness of hospital admission.[7]

Second, as stated previously, HCFA contracts with SuperPRO to rereview a random sample of medical records initially reviewed by PROs. Approximately 400 records per PRO are reviewed in each SuperPRO review cycle, depending on the expected overall PRO review workload. Individual SuperPRO review cycles during the second round of PRO contracts covered roughly six months of PRO review activities. Data from this source are referenced as SuperPRO data. One of the most important characteristics of this data source is that the SuperPRO review process seeks to replicate that of the individual PRO in most respects. We obtained data directly from SuperPRO on the results of these rereviews, as well as the results of the initial PRO review of the same cases.[8] As a result, we present information not only on differences between the final SuperPRO determination of appropriateness and the initial PRO decision but also on the differences in the rate at which PRO and SuperPRO nurses refer cases for physician review.

Review of Medical Records

We contracted with Value Health Sciences, Inc., to conduct reviews of 246 cases of elective coronary angiography that had been previously reviewed by PROs across the country, as well as by SuperPRO.[9] All elective coronary angiographies with discharge dates between December 15, 1986, and July 8, 1987, on which SuperPRO had completed their reviews were selected.[10] The cases were reviewed using Rand-style protocols that

[6]The HCFA data summary that we worked with contained accepted reports through February 1989. The percentage of PRO reports accepted to reports required varied from 62 percent to 100 percent, with 50 PROs being at least 90 percent complete.

[7]The rates reported are for cases initially denied. The final rates—after denials reversed based on reconsiderations and appeals are removed—would be somewhat lower.

[8]The data from SuperPRO presented in this report are limited to roughly the same time period as that of the HCFA data we present.

[9]Value Health Sciences, Inc., is a private sector firm that specializes in UR products that incorporate scientific research findings. The firm's senior vice president is Mark Chassin, M.D., who was one of the researchers involved with the Rand work on appropriateness indications.

[10]Since the cases were selected from the files of SuperPRO, they are not formally representative of the practice of coronary angiography in the nation. However, the cases do come from all regions of the country.

had been updated to reflect appropriate medical practice as of 1987.[11] The results of these reviews provide a more current estimate of the rate of potentially inappropriate coronary angiography in the Medicare population and serve as a replication of the earlier Rand study.

Report Organization

Because an understanding of the UR process is essential to interpreting the evidence concerning the levels of inappropriate care in the Medicare program, we outline in chapter 2 the basic approaches to UR that are available and describe the procedures actually used by private sector organizations and by PROs. We then discuss the advantages and disadvantages of the different approaches in reducing the inappropriate use of hospital services. We also discuss the limited information that is available on the cost-effectiveness of UR programs.

In chapter 3, we examine available estimates of the level of inappropriate care in the Medicare program. Whenever possible, we contrast estimates derived from PRO review of Medicare claims with those derived from other sources, in order to evaluate the extent to which PROs are discovering inappropriate care. We explore the differences between the PRO review process and those processes used by other review organizations as potential causes of differences in the rates at which inappropriate care was found.

In chapter 4, we pull together the different lines of evidence presented in chapters 2 and 3 and offer suggestions whose adoption, we believe, might result in more effective methods for dealing with inappropriate care in the Medicare program. These suggestions are based both on approaches used in the private sector that could be effectively applied to the Medicare program and on aspects of the current PRO program that may lessen the ability of individual PROs to identify inappropriate care.

[11] M. Chassin, et al., Indications for Selected Medical and Surgical Procedures—A Literature Review and Ratings of Appropriateness: Coronary Angiography (Santa Monica, Calif.: The Rand Corporation, 1986).

Reviewing Medical Care for Appropriateness: Medicare and Private Sector Approaches to Utilization Review

Introduction

In this chapter, we will describe the UR activities performed by a variety of private sector UR organizations and by PROs.[1] Any organization that chooses to review the appropriateness of health care must make many important decisions about how to operationalize the concept of appropriateness and how to design the UR system in which judgments about the appropriateness of health care are made. Many options are available, and the choice of a particular option may influence the program's effectiveness. For example, decisions must be made about when to conduct the reviews, what aspects of appropriateness to review, what process to use to identify instances of inappropriate care, what review protocols to apply, what cases to review, and what to do about any inappropriate care discovered. These decisions provide the framework for our examination of UR programs. In addition to describing private sector and PRO approaches to UR, we will also discuss their relative advantages and disadvantages when these approaches differ. The information contained in this chapter was obtained from the published literature and supplemented by our interviews with individuals knowledgeable about specific aspects of UR.

Although the decisions concerning UR will be discussed in sequence, they are interrelated; any individual decision may have implications for several others. Of critical importance is the question of when review should be conducted—before, during, or after a period of hospitalization. A decision on the timing of review influences the types and sources of information on which an appropriateness decision may be based, as well as the actions that may be taken in response to finding that an actual or proposed health care service is inappropriate. Another critical decision is the review protocol (that is, criteria) to be used in reaching an appropriateness determination. Whether the protocol should focus on patients with specified diagnoses or be independent of diagnosis is of concern. The balance between criteria that are general and inclusive and those that are specific and mutually exclusive has important implications for the validity and reliability of review decisions. UR approaches are still evolving, and there currently is little systematic knowledge about the operational strengths and limitations of particular options or, for that matter, the overall effectiveness of particular UR programs.

[1] For a more extensive discussion of all PRO responsibilities, see U. S. General Accounting Office, Medicare: Improving Quality of Care Assessment and Assurance, GAO/PEMD-88-10(Washington, D.C.: May 1988).

The Timing of Reviews

Available Options

One of the most important decisions in setting up a UR program is the timing of reviews. The reviews of appropriateness can be conducted prospectively, concurrently, and retrospectively. Prospective reviews occur before the medical services are provided and are typically based on information regarding the patient's medical needs given to the UR program by the attending physician. Information may also be obtained from the patient about the frequency and duration of specific symptoms. The reviews may focus on any (or all) of the appropriateness issues, including the need for the planned medical care, the need for providing that care in the hospital, and, if hospitalization is needed, the expected length of stay.

A variety of terms have been used to describe programs of prospective review. Preadmission, precertification, or prior approval programs refer generically to prospective reviews of proposed hospital admissions, regardless of the appropriateness issues being reviewed. Preprocedure review refers to the review of the appropriateness of a proposed surgical procedure either before or after hospital admission, but prior to surgery. Second surgical opinions refer to the prospective review of a proposed treatment plan by a second, independent physician.

Concurrent review occurs while the medical care is being provided. Typically, concurrent review focuses on whether a patient needs to remain in a hospital and on discharge planning (that is, when is the patient ready for discharge, where should the patient go after discharge from the hospital, and what services need to be arranged).

Retrospective review occurs after the patient has been discharged from the hospital. One approach is to conduct manual reviews of the actual medical records of individual patients to identify possibly inappropriate care. A second approach is to use computer analyses of information taken from hospital bills or abstracted from medical records to develop statistical norms of utilization in a particular area and to identify particular providers whose patterns of utilization are significantly different from those norms. This is called profiling. Like prospective review, retrospective UR can be used to address either of the appropriateness questions.

One advantage to both prospective and concurrent reviews is the potential to identify and rule out potentially inappropriate care prior to delivery. This reduces the beneficiary's exposure to potentially dangerous treatments and eliminates the need to deny payment after resources have been committed. However, prospective review may be more intrusive on the individual physician's decision-making function. Retrospective review interferes less with the physician's practice of medicine but places him or her (and the hospital) at financial and professional risk for having provided inappropriate care. It also means that providers have powerful financial and professional incentives for justifying their actions.

Private Sector Programs

Private sector organizations typically rely on prospective reviews of appropriateness at the individual patient level and generally focus on all elective admissions. For example, a 1987 survey of preferred provider organizations found that 92 percent had prospective review programs. Similarly, 95 percent of the benefit managers of 100 large corporations participating in a 1988 survey reported conducting prospective review of hospital admissions. In a 1987 study, 55 of 62 responding Blue Cross and Blue Shield plans conducted prospective reviews. Further, these surveys show that most organizations conducting prospective reviews also conduct concurrent reviews.

Because of the strong emphasis on prospective review of individual cases in the private sector, retrospective reviews of individual cases tend to consist primarily of audits of particularly costly hospital stays or suspect claims. There are also medical record audits to confirm the information provided in the prospective reviews. In addition, many organizations use retrospective profiling of claims information to identify providers whose utilization patterns differ significantly from the norm in a particular geographical area. The identification is based on unusual admission patterns, lengths of stay, and resource utilization. Providers associated with these aberrant patterns may then be subjected to intensified review, either prospectively or retrospectively. These reviews are particularly important because much private sector reimbursement to hospitals and physicians is still on a fee-for-service basis. Further, some organizations are beginning to use this information in selecting providers with whom to negotiate contracts for particular services.

PRO Program

The admission and invasive procedure reviews, which are the primary PRO activities that focus on appropriateness, are both retrospective reviews. Admission review is intended to detect inappropriate admissions. The invasive procedure review requires PROs to review, for appropriateness, all surgical procedures and any other procedure that would affect DRG assignment.[2]

PROs are required by the Third Scope of Work to engage in extensive profiling activities to identify providers for further, in-depth review and to monitor local practice patterns, including any changes in them that might be associated with PRO review activities. PROs are, however, not required to conduct any specific appropriateness reviews based on profiles. While HCFA's evaluation of PRO performance, called PROMPTS, contains a section related to the activity of profiling, none of the elements evaluates the extent to which PRO uses its profiling activities to target or focus appropriateness reviews.[3] Finally, while PROs are not required to maintain paper copies of all their profiles, they are required to produce them on request for either HCFA or a local provider.

PROs also do prospective (that is, preadmission/preprocedure) reviews of selected surgical procedures; under the Third Scope of Work contracts, these reviews are limited to ten procedures.[4] The review of two procedures—cataract extractions and carotid endarterectomies—is mandated. Each PRO selects eight other procedures from a list of candidate procedures provided by HCFA and then provides empirical justification to HCFA for its choices.[5] In addition, PROs must approve all cases where the use of an assistant surgeon is proposed for cataract procedures. These prospective reviews are intended to consider both the appropriateness of the proposed procedure and the appropriateness of admission (if inpatient treatment is proposed).

[2]The invasive procedure review was not added to the PROs' scope of work until the third round of contracts. As a result, there were no data on the results of these reviews available when we conducted our study.

[3]PROMPTS is an acronym for Peer Review Organization Monitoring Protocol and Tracking System.

[4]Under the second round of PRO contracts, only five procedures were subject to preadmission review. The information presented here reflects the current scope of work for the third round of PRO contracts. The Congress mandated that PROs review ten procedures in conjunction with an as yet uninitiated second opinion program.

[5]The list provided by HCFA in the third PRO Scope of Work contains 11 procedures: cholecystectomy, major joint replacement, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty, laminectomy, complex peripheral revascularization, hysterectomy, bunionectomy, inguinal hernia repair, prostatectomy, and pacemaker insertion. PROs may also choose other procedures not on HCFA's list with adequate documentation of the need to review those procedures and HCFA's approval.

PROs are not required to engage in concurrent review. In part, this is because hospitals under Medicare Prospective Payment System (PPS) already have strong incentives to minimize length of stay. However, PROs must investigate cases in which patients request a review of physicians' decisions to discharge them.

Discussion

Private sector UR programs generally place greater emphasis on prospective and concurrent reviews of appropriateness in individual cases than do PROs. The PRO emphasis reflects, in part, the fact that PPS makes the review of the length of an individual hospital stay less important for the Medicare program than it is in the private sector. In addition, in the Medicare program, if a hospital stay is determined to have been inappropriate, the patient will not generally incur any financial liability for the hospital stay. In the private sector, the hospital can attempt to collect any difference between the amount charged to the patient and the amount paid by the payor. As a result, in the private sector, prospective review may protect patients from unexpected bills.

Both the private sector and PROs engage in extensive retrospective profiling of claims data. However, in the PRO program, the resources required to conduct individual retrospective case reviews, combined with the lack of a clear incentive to use profiles to guide review activities, may make profiling less effective than it might otherwise be.

The Appropriateness Issues Addressed

Available Options

As previously noted, there are two major questions concerning appropriateness that may be addressed in UR programs. First, are the proposed or delivered medical services needed by the individual patient? Second, where should the needed services be (or have been) provided, and how long should treatment last? The first question focuses on the needs of an individual patient; the second, on the proper level of care for meeting those needs—hospital, ambulatory surgery center, or other alternative.

Private Sector Programs

Second opinion programs, both voluntary and mandatory, are frequently used by the private sector to address the appropriateness of

individual medical services. The cases reviewed typically involve elective surgical procedures. These programs permit (in the case of a voluntary program) or require (in the case of a mandatory program) a patient for whom surgery is planned to obtain an opinion on the need for surgery from a second physician.

Almost all private sector programs review the appropriateness of hospital admissions prospectively as a part of preadmission certification programs. In determining the proper level of care, a large number refer to lists of surgical procedures approved for delivery in outpatient settings, such as hospital outpatient surgery departments or ambulatory surgical centers, to attempt to divert potential hospital admissions to alternative sources of care. An attempt to control length of stay is also a large component of most private sector programs.

PRO Program

As previously noted, PROs are currently required to conduct invasive procedure and admission reviews on each case selected for retrospective review. Length-of-stay is typically not reviewed because, under PPS, hospitals are reimbursed a set amount for each admission, regardless of the length of stay.[6] Prospective reviews of the ten PRO-designated procedures involve a consideration of whether the procedure is needed and, if so, where it should be provided. In the case of procedures that are always done in the hospital, the determination that the procedure is appropriate implies that the admission is also appropriate. For procedures that can be done under certain circumstances in an outpatient setting, a review of the need for hospital admission is required.[7]

Discussion

While the private sector and PROs differ in regard to when they examine the appropriateness of hospital admissions, they both actively review this aspect of hospital care. However, their interest in reviewing the appropriateness of individual medical services has been limited by the relative unavailability of reliable and valid review criteria. The persuasiveness of recently published research on the extent of inappropriate use of particular procedures, as well as growing interest in this problem in both the private sector and Medicare, has spawned a new thrust

[6]The only exceptions to this policy are made for cases, termed outliers, that greatly exceed the expected length of stay for a particular DRG, and for specialty hospitals. Hospitals may receive more than the established DRG payment for outlier cases.

[7]Only three of the candidate procedures are potential outpatient procedures: cataract extraction, bunionectomy, and hernia repair.

within the medical profession and health policy community toward the development of review protocols (also referred to as practice guidelines, standards, or practice parameters) to determine the appropriateness of individual medical procedures.

Evidence of this increased interest can be seen in the number of meetings devoted to the issue of "practice guidelines." In September 1987, the Council of Medical Specialty Societies held a conference on experiences of, and risks associated with, standard setting. In September 1988, the Subcommittee on Health and the Environment of the House Committee on Energy and Commerce held a hearing on the development and use of medical practice guidelines. Shortly afterwards, in October 1988, the Physician Payment Review Commission also held a conference as part of their effort to develop, by early 1989, a policy position on the use of such guidelines. At those meetings, statements by representatives of various groups, including the American Medical Association and Council of Medical Specialty Societies, clearly indicated that they recognized the pressure to move forward in the area of practice guidelines. Once such guidelines become available, PROs and private review organizations can be expected to expand their reviews of the appropriateness of individual medical services.

The Process of Identifying Inappropriate Care

Available Options

The process of identifying individual instances of inappropriate care generally involves using an explicit protocol to examine information about specific patients and select questionable cases for further review.[8] This initial "screening" can be either a manual or an automated process. Manual screening generally begins with a nurse or other trained medical professional evaluating the available medical information (for example, the medical record) for an individual patient using an explicit protocol or protocols. Automated screening involves entering information about a case into a computer and then using a computerized protocol to make the initial screening determination. If the medical care is judged to be

[8]See page 24 for a definition of the term "protocol" and a description of how protocols are used.

appropriate, the case is approved. If not, it is referred to a physician for further review and a final determination.

When cases are referred to a physician for further review, the physician's determination may be based on his or her own knowledge, training, and experience rather than on an explicit protocol. Systems differ in the extent to which individual physicians are expected to conform to a particular set of medical practice norms. For example, in health maintenance organizations, individual physicians may be expected to adhere to the organization's norms for the use of specific procedures or settings. Systems may also differ in the extent to which additional information justifying the admission is sought from the attending physician prior to a final determination. Finally, depending on the review system, the physician's final determination of appropriateness may also be appealed to higher levels.

Retrospective reviews of individual cases are typically based on medical records. At a minimum, PROs are required to examine the face sheet, the physician's attestation statement, the physician's admission note, the discharge summary, the history and physical, physicians' progress notes, and physicians' orders. Other parts of the medical record—for example, lab reports, pathology reports, and nurses' notes—may be requested and reviewed as necessary. When the medical record is reviewed at the hospital (that is, "on-site"), all of the pertinent information on which to base a judgment should be available. However, when the medical record is copied and reviewed off-site, the photocopies of records may not be entirely complete or legible, and some may therefore not include all the information the reviewers might want or need, although reviewers can request that missing information be provided.

In prospective reviews, the information available to the reviewers is typically much more limited. The reviewer will usually have only information from the attending physician concerning the proposed treatment and why it is necessary. This information could include the results of tests that indicate the need for treatment. In some cases, the reviewer may also contact the patient to obtain additional information.

Private Sector and PRO Programs

Both PROs and the private sector UR programs use primarily the same, manual approach to reviewing individual cases. Regardless of whether the review is prospective or retrospective, much of the information is initially examined by a trained medical reviewer and a decision then made on the basis of written protocols. In the case of initial PRO review,

the medical record is most often reviewed on-site; final physician review is more typically done off-site. In both settings, if a case is referred to a UR physician for further review, the UR physician may interact personally (for example, on the telephone or through correspondence) with the attending physician.

However, both PROs and private sector UR programs are considering automating their initial screening processes. HCFA is currently working on a project to computerize much of the PROs' retrospective screening of individual cases. In addition, in some of the UR firms that do a large volume of preadmission certifications by telephone, the information may already be fed directly into a computer and certain aspects of the review guided by protocols that automatically compare the information to a protocol programmed into the computer.

Discussion

There are few, if any, important differences in the processes by which the private sector and PROs conduct their reviews of individual cases. The differences are primarily in the timing of the review and the source and extent of the information available to the reviewer. In private sector prospective reviews, the information must come directly from the physician and/or patient and is necessarily limited. In PRO retrospective reviews, the entire hospital medical record may be examined and additional information requested from the attending physician if it is not included in the medical record.

Protocols Used to Screen Cases

Available Options

In this report, the term "protocol" refers generically to a tool used to guide decision making about the appropriateness of medical care. In theory, protocols could be used to definitively determine the appropriateness of care. In practice, protocols are used in UR programs to screen or identify cases of questionable care for subsequent review by physicians.[9]

[9]Appendix I contains a more detailed discussion of available protocols and their use by various UR organizations.

Screening protocols can be divided into those that are applied to patients with specific diagnoses or who are undergoing specific procedures and those that are independent of diagnosis. Diagnosis-specific protocols indicate, for a given medical problem, which diagnostic and therapeutic services are appropriate. We include in this group procedure-specific protocols that list the indications and contraindications of a particular diagnostic or therapeutic intervention. Diagnosis-independent protocols incorporate criteria that are relevant to a broad range of patients, regardless of their diagnosis, and for whom the severity of illness, signs, and symptoms may determine the appropriateness of care.

Both diagnostic-specific and diagnostic-independent protocols are explicit—that is, they are written, detailed, and used in a standardized and consistent manner over time. Explicit criteria may be differentiated from implicit criteria, which typically are not written and reflect the reviewer's professional training, experience, and judgment.[10] The primary advantage of the implicit approach is that it allows the reviewer to consider all relevant factors affecting the circumstances of the patient and the care provided, not all of which may be reflected in explicit criteria. Implicit physician judgments can be said to have face validity because most observers believe that physicians, as medical care professionals, are in the best position to make judgments about appropriateness. The primary advantage of explicit protocols is that, because they are written, they can be used by less highly trained individuals. Thus, they are cheaper to apply. In addition, because they are standardized, they may produce more reliable results, in the sense of obtaining the same or similar results on multiple reviews or when using different reviewers.

Private Sector Programs

It is difficult either to assess the totality of protocol-development activity or to describe the protocols actually used in ongoing private sector UR programs. There is no common registry for such efforts, and many of the protocols developed or used by commercial firms are proprietary. Nevertheless, in the course of our study, we had the opportunity to ask many different people in a variety of organizations about their use of protocols.

[10]A protocol can, in fact, vary along a continuum from very explicit to totally implicit. A protocol toward the implicit end of the continuum might simply specify the types of information that should be considered in making a determination. A very explicit protocol might specify that a particular value on a particular diagnostic test is essential in order to determine that a procedure is appropriate.

Many commercial developers of UR products, as well as firms conducting reviews under contract, have developed their own diagnostic-specific protocols with the assistance of physician consultants as well as by drawing on protocols developed by others, including the medical specialty societies. For example, InterQual, a developer of UR products, has compiled a manual, SIM III: Surgical Indications Monitoring, that contains over 140 pages of protocols for specific surgical procedures. Publications dealing with UR in the private sector, as well as a number of the people we interviewed, indicated that more and more organizations will be developing and using diagnosis-specific protocols in the near future.

As is the case with diagnostic-specific protocols, there is no systematic information on what diagnosis-independent protocols are used in the private sector. One observer told us that firms that specialize in UR usually use a review protocol like the Appropriateness Evaluation Protocol (AEP). AEP emphasizes the appropriate level of care regardless of the diagnosis. Another individual we interviewed estimated that one third of the organizations use AEP; one third use the Intensity of Service, Severity of Illness, Discharge Screens, and Appropriateness review system (ISD-A) developed by InterQual; and one third use protocols developed by their own staff.

PRO Program

PROs are contractually required by the Scope of Work to use written criteria (that is, protocols) in their screening activities. Each PRO is responsible for developing its own protocols in consultation with physicians in the state who are actively engaged in the practice of medicine. The protocols are subject to HCFA review, must be sent to all providers and physician organizations in the state, and must be available to anyone on request.

In 1988, we conducted a survey of PROs to obtain information on a variety of topics, including organizational structure, resources, staffing, and review processes.[11] One of the questions asked PROs what protocols they used in reviewing care for potential utilization problems. Of the 53 PROs responding to the survey, many reported using locally developed diagnostic-specific protocols or HCFA's coverage criteria. Four PROs reported

[11]Some of this information has been presented in a report to the Subcommittee on Health of the House Committee on Ways and Means. See U.S. General Accounting Office, Medicare PROs: Extreme Variation in Organizational Structure and Activities, GAO/PEMD-89-7FS (Washington, D.C.: November 1988).

using the previously mentioned SIM-III. In terms of diagnosis-independent protocols, most PROs (39 of 53) used the ISD-A protocol or some variant of it. An additional 7 PROs reported using AEP.

Discussion

While the exact contents of many screening protocols are not publicly available, our interviews suggest that many of the protocols used by PROs and the private sector are either developed internally by the individual program for its own use or adapted from other commercial or public protocols. As a result, few of the protocols are likely to have been systematically tested for reliability and validity. In addition, the use of many different protocols means that there may be variation in the types of cases found to be appropriate. Finally, the growing interest in the development of national practice guidelines may lead to a decline in the use of relatively untested local protocols in the future.

The Types of Cases Reviewed

Available Options

Unlike many of the issues discussed previously, the question of what cases to review does revolve around cost considerations. That is, reviewing all cases would ensure that no potential cases of inappropriate care went unexamined. However, the resources required to conduct reviews of all cases must be weighed against the potential for identifying cases of inappropriate care. Targeting review on cases with a greater potential for involving inappropriate care is one way to tilt the balance. However, this approach can be effective only if a successful strategy for targeting is available. One approach might be to target procedures that can be done safely and more cheaply in an outpatient setting than in a hospital. Another approach would be to target admissions for specific diagnoses that have accounted for significant numbers of inappropriate admissions or surgical procedures in the past.

Private Sector Programs

Given the large number of private sector UR programs, there is considerable variation in the types of cases that are targeted for review. In many benefit plans, all elective hospital admissions are required to have

preadmission certification.[12] Other plans focus on elective surgeries that could be performed on an outpatient rather than inpatient basis or on specific diagnoses for which inappropriate utilization is likely, thereby attempting to maximize the return on investment in UR. With the likely increase in reviews directed at the appropriateness of individual procedures, the targeting of these programs may well change in the future.

PRO Program

The categories of cases that PROs review are specified to a large extent by HCFA and reflect the variety of discrete review tasks assigned to PROs. Table 2.1 presents the categories of cases that PROs are required to review retrospectively (under the second round of PRO contracts), the percentage of the total number of reviews that each category represents, and the admission denial rate for that category of case.[13] The denial rate associated with the random sample of cases is 2.2 percent. Over 60 percent of all other reviews occur in categories of cases with admission denial rates that are below that associated with the random sample cases. However, cases referred to PROs by either hospitals or fiscal intermediaries specifically because the appropriateness of admission is already being questioned have relatively high admission denial rates. In addition, when PROs conduct intensified review of particular providers based on identified problems (that is, targeted review), they find an increased rate of inappropriate admissions. These results suggest both that PRO reviews are not well targeted to identify inappropriate admissions and that better targeting is feasible.

[12]Admission urgency is generally placed on one of three levels. Emergency admissions cannot be delayed. Urgent admissions can be delayed 24 hours or more. Elective admissions can be scheduled at the convenience of the patient or physician.

[13]In this report, admission denial rate refers to that percentage of cases reviewed in which the admission was determined to be inappropriate.

Table 2.1: Admission Denial Rates for Categories of Cases Reviewed Retrospectively by PROs[a]

Type of case	No. of reviews	Percent of total reviews[a]	Admission denial rate
Readmissions[b]	2,221,036	33.2%	1.9%
Utilization objectives[c]	1,171,137	17.5	3.5
Transfers[d]	885,163	13.2	1.4
Random sample[e]	683,701	10.2	2.2
DRGS 468, 462, 088	532,526	8.0	2.1
Intensified review[f]	484,910	7.2	3.9
Day and cost outliers	460,620	6.9	0.8
Specialty hospitals	131,093	2.0	2.3
Hospital notices of noncoverage	69,593	1.0	7.5
Fiscal intermediary referrals[g]	25,293	.4	8.8
Other cases	34,107	.5	1.4
Overall	6,699,179	100.1%	2.3%

[a]PRO data reflecting reviews done under the second round of contracts as collected and tabulated by HCFA. Total does not equal 100 percent due to rounding.

[b]A readmission, under the second round of contracts, is any case in which a patient is readmitted to a PPS hospital within 15 days. Each review may involve the review of two or more admissions.

[c]PROs were required to establish specific utilization objectives under the second round of contracts and to review specific cases to demonstrate that they met their objectives. Such objectives are optional under the third round of contracts.

[d]This category of cases includes transfers from a PPS hospital to another hospital, exempt unit, or "swing bed" (that is, skilled nursing home-type bed in certain types of hospital). Each review may involve the review of two or more admissions.

[e]A 3 percent random sample of each hospital's discharges.

[f]Intensified review of particular hospitals based on findings from reviews of all cases.

[g]The fiscal intermediary referrals include cases identified by the Medicare Code Editor, potentially non-covered admissions, and other cases referred by the fiscal intermediaries.

In terms of preadmission reviews, many PROs chose to review largely inpatient surgical procedures under the second round of work. Because inpatient surgical procedures by definition require hospital admission, reviews of admission necessity would not be expected to result in the denial of many hospital admissions unless the need for the procedure itself was questioned.[14] As mentioned previously, the procedures specified by HCFA for PRO review in the third round of PRO contracts are again largely inpatient procedures. Unless PRO reviews for the appropriateness of the surgical procedure itself result in increased denials, the situation is not likely to change much.

[14]Data on admission denials under PRO prospective review are presented in chapter 3.

Discussion

Increasing the cost-effectiveness of appropriateness reviews depends, in part, on targeting review on the types of case where inappropriate use is more likely to occur. The private sector attempts to do this in its prospective reviews by focusing on specific diagnoses, elective surgical procedures that can be performed on an outpatient basis, or more generally, elective hospital admissions. Similarly, PROs' limited prospective reviews are focused primarily on inpatient procedures that research and anecdotal information suggest may be used inappropriately. (See chapter 3 for a discussion of the literature on the inappropriate use of particular procedures.)

However, the data just presented suggest that PRO retrospective reviews are focused on categories of cases in which the admission denial rates are generally lower than for a randomly selected group of cases. Given the variety of different review responsibilities assigned to PROs, this selection of cases may be quite appropriate. However, if controlling the inappropriate use of hospital services is the goal, the selection of cases could be better targeted.

Data from a number of studies of Medicare hospital admissions indicate that many of the cases that are inappropriate hospital admissions could have been treated on an outpatient basis. For example, of the Medicare admissions judged to be inappropriate by SuperPRO physicians, 65 percent should have received treatment in either the attending physician's office or an ambulatory surgery center. A study by the Department of Health and Human Services Inspector General found that 89 percent of the inappropriate Medicare admissions should have been treated as outpatients.[15] The specific findings in this study include the following:

- surgical admissions had a higher inappropriate rate (11.8 percent) than did medical admissions (9.6 percent);
- elective admissions had a much higher rate of inappropriate admissions (19.7 percent) than did either urgent (8.3 percent) or emergency admissions (2.4 percent);
- lens procedures was the DRG with the highest rate of inappropriate admissions (80.2 percent), with a group of seven DRGs (lens procedures; pathological fractures; esophagitis, gastroenteritis, and miscellaneous digestive disorders; otitis media and upper respiratory infections; medical back problems; respiratory neoplasms; and diabetes) accounting for

[15]Health Data Institute, National DRG Validation Study (Lexington, Mass.: Health Data Institute, 1987).

24.2 percent of all inappropriate admissions (and, with the exception of lens procedures, all of these being medical DRGs);
- among reasons for admission, planned procedures had the highest inappropriate rate (14.6 percent);
- for principal procedures other than lens procedures, endoscopies and biopsies accounted for relatively high rates of inappropriate admissions (11.1 to 23.4 percent, depending on the procedure).

In a study that focused on the issue of targeting UR, Payne analyzed 7301 Medicare hospital discharges in the 65 highest volume DRGs.[16] Based on ratings using AEP, medical DRGs such as medical back problems and degenerative nervous system disorders had higher overall rates of inappropriate admissions than did surgical DRGs.[17] Individual medical DRGs with the highest rates of inappropriate admissions included medical back problems, bone diseases, and pathological fractures and musculoskeletal and connective tissue malignancy. Surgical DRGs with high rates of inappropriate admissions included coronary artery bypass surgery without cardiac catheterization, major chest procedures, and major small and large bowel procedures. Focusing PRO reviews on these types of cases, which are frequently associated with inappropriate admissions, might improve review effectiveness.

Alternative Responses to Inappropriate Care

Available Options

At some point in the process of reviewing hospital care for appropriateness (either retrospectively or prospectively), a determination is made that the medical care is either appropriate or inappropriate. For care that is deemed appropriate, the result is clear. The medical care takes (or has taken) place, and payment follows. For care that is deemed to be

[16]S.M.C. Payne, "Targeting Utilization Review to Diagnostic Categories," Quality Review Bulletin (December 1987), pp. 394-404. (The cases came from three previous studies of inappropriate hospital admissions using AEP.)

[17]One potential reason given in the article for the results for medical DRGs relative to surgical DRGs is that the basic AEP is sensitive to elective admissions in which no hospital-level services are provided within the first 24 hours but does not identify any surgical admissions that might potentially have been outpatient surgery.

inappropriate, some additional action is required. The number of potential alternatives is large, but all of them would likely involve some limitation on the amount of money that the payor will reimburse for the inappropriate care.

If the only option is to deny all payment, this creates powerful incentives for both the reviewers and those they are reviewing to resort to alternative, informal approaches to dealing with the situation. For example, the reviewers may talk to the attending physician, gather additional information, and develop a rationale for approving the admission, thereby avoiding the need to confront a professional colleague overtly. At the same time, the attending physician would be informed that the appropriateness of the admission was questionable, in hopes that his or her behavior would change in the future. Alternatives to outright denial of payment, such as questioning the proposed treatment and suggesting other modes of therapy, typically allow a more collegial, less confrontational approach to be taken.

Private Sector Programs

In private UR organizations, the final decision in a prospective review is almost always "advisory." That is, the decision is that the admission cannot be certified as appropriate. It is not a final decision to deny payment if the admission should, in fact, occur. The final payment decision is usually deferred to a later point in time. As a result, the determination that the admission is unnecessary leaves the patient and physician with a decision as to how to proceed. They may decide not to carry out the treatment plan, to use an alternative setting or approach, or to go ahead with the hospital admission.

Private sector organizations use a variety of approaches to persuade patients and physicians to comply with the certification decisions, short of totally denying payment for uncertified care. First of all, the review nurse and UR physician try to avoid having to deny certification by persuading the physician to adopt an alternative strategy for dealing with the particular patient. Depending on the situation, this might involve suggesting either appropriate alternatives to inpatient care (such as ambulatory surgery) or alternative treatment strategies that would be appropriate for the particular patient. To the extent that this strategy of suggesting alternative approaches is successful, it helps hold down denial rates.

Second, benefit plans are structured to encourage compliance with UR decisions. Instead of having his or her payment denied for noncertified

hospital use, the patient may be required to pay increased copayments or deductibles. For example, employees of one company have their coverage reduced from 80-100 percent coverage to 50 percent if they do not comply with the UR program. Other employees receive 100 percent coverage if they comply with the UR program but must pay $250 if they do not comply.

Third, some plans require employees to obtain second opinions as a way of providing them with additional information on the appropriateness of a particular treatment plan. If the second opinion differs from the first, the additional information provided should allow them to make a better informed choice about the appropriate treatment for them. Then, in addition to having the second opinion paid for, the patient is free to choose whatever treatment plan he or she prefers, without monetary penalty. As experience is gained with second-opinion programs, some companies are restructuring them to require that second opinions be obtained only from carefully preselected physician panels.

PRO Program

Formally, in PRO reviews of appropriateness, once a decision is made that the care was inappropriate, a denial of payment automatically follows.[18] The denial involves both the hospital DRG payment and the physician's reimbursement.[19] The providers have the right to request a reconsideration of the decision and also to appeal administratively PRO determinations.

Informally, our interviews with officials of three PROs indicated that PRO physicians often attempt to deal with cases of questionable appropriateness in the kind of nonconfrontational, collegial style described previously. They often have extensive interactions with the attending physicians to discuss particular cases and to discover if there is any additional information that might justify the admission in question.

[18]Early in the implementation of PPS, PROs could allow payments to hospitals for inappropriate admissions under a waiver of liability that required a "favorable presumption" that the hospital did not know that the admission was inappropriate. HCFA officials indicated to us that few, if any, cases are now paid under waivers of liability. HCFA argues that all hospitals should now know which admissions are appropriate and which are likely to be inappropriate and no longer paid for under waivers of liability.

[19]In most cases, PRO notifies the fiscal intermediary to recover any payment made to the hospital. The fiscal intermediary is responsible for notifying the carrier that the hospital admission was inappropriate. The carrier is responsible for making its own determination concerning whether to deny payment to the physician for any services provided during that admission. There are two exceptions when PRO notifies the carrier directly and the physician payment is automatically denied: 1) cases involving inappropriate surgical procedures, and 2) cases involving physician charges associated with denied cost-outlier services.

Cost-Effectiveness of Utilization Review

The limited information available suggests that both the private sector and PRO UR programs are cost-effective. That is, both save more money on health care expenditures than is spent in conducting the reviews. It has been estimated that PROs deny $2.29 in payments for each $1.00 spent on review (that is, PROs have a cost-effectiveness ratio of 2.29:1). Similar claims have been made regarding the cost-effectiveness of utilization review in the private sector. For example, Ermann cites studies of private sector UR programs with cost-effectiveness estimates ranging from 3:1 to 30:1.[20] However, as Ermann notes, few of the estimates seem to be based on scientifically sound evidence. As a result, claimed savings must be carefully assessed.

In one of the few empirical studies of the cost-effectiveness of utilization review, Feldstein and others analyzed the claims' experience over a two-year period of 88 insured private sector groups with at least one UR activity and 134 groups without a UR program.[21] The UR activities involved were preadmission certification by telephone, preadmission certification and continued stay review by medical personnel at the hospital, and concurrent review by medical personnel at the hospital of the treatment plan for the appropriateness of the length of stay and use of ancillary services (for example, laboratory tests or physical therapy). The groups with at least one UR activity had a reduction in admissions of 12.3 percent and in expenditures of 11.9 percent relative to the groups without a UR activity. The cost-effectiveness ratios for all groups with utilization review was 8.7:1; in groups with high prior utilization, the ratio was 28.3:1.

Summary and Conclusions

Overall, we found that PROs and private sector UR programs are fairly similar in terms of the appropriateness issues addressed and the process used to identify inappropriate care. Both have focused on identifying inappropriate admissions but are becoming more involved in reviewing the appropriateness of individual services. Both employ medical professionals using explicit protocols to screen cases of potentially inappropriate care for final determinations by physicians. The types of screening protocols used appeared to be similar, although the information available to us on this issue was limited.

[20]D. Ermann, "Hospital Utilization Review: Past Experience, Future Directions," Journal of Health Politics, Policy, and Law, 13:4(1988), pp. 683-704.

[21]The UR activities were conducted by HealthCare Compare, a private sector UR company for a private health insurance company. See P. Feldstein, T. Wickizer, and J. Wheeler, "The Effects of Utilization Review Programs on Health Care Use and Expenditures," New England Journal of Medicine, 318(1988), pp. 1310-14.

There are significant differences, however, in how PROs and private sector UR programs have addressed three important issues in structuring their review programs. First, the private sector depends heavily on prospective review of elective hospital admissions. Retrospective reviews are used to identify billing inconsistencies and to profile practice patterns of individual providers. By contrast, most of the review resources in the PRO program are devoted to retrospective reviews. A total of only ten surgical procedures are subject to PRO prospective review, and most of those procedures are generally performed on an inpatient basis.

Second, PRO retrospective reviews of many of the categories of cases mandated by HCFA do not yield higher rates of admission denial than does review of a simple random sample of cases. While the current selection of cases may be necessary to accomplish other PRO objectives, it lessens the likelihood of PROs identifying inappropriate admissions or procedures. Private sector UR programs focus on either all elective hospital admissions or targeted groupings of potentially inappropriate admissions (such as potential outpatient surgeries). The evidence concerning inappropriate Medicare admissions generated by researchers suggests that the types of Medicare cases found to have involved inappropriate admissions are quite similar to those targeted by the private sector for prospective review, but that they are not currently singled out for special attention by PRO reviewers.

Finally, the determination that an admission is inappropriate has different implications in the private sector and Medicare programs. In most cases, a prospective "denial" of certification in the private sector is actually a warning to the provider that the admission is questionable; however, an individual and his or her physician can choose to disregard the decision. (The consequences of going ahead with the admission vary, depending on the health benefit plan.) In the Medicare program, a determination that an admission is inappropriate results in denial of payment. The punitive aspects of retrospective denials of payment and the lack of formal options for dealing with inappropriate care (other than denial) may result in an environment in which it is difficult for PRO physicians to both maintain good working relationships with the local medical community and assure that the Medicare program is paying only for appropriate care.

Levels of Inappropriate Care in the Medicare Program

Introduction

One impetus for the Subcommittee's question about the level of inappropriate care in the Medicare program came from an informal comparison of the relatively low rates of denial for inappropriate admissions reported by PROs with published accounts of considerably higher rates of inappropriate use of specific surgical and diagnostic procedures in the Medicare population.[1] Although this informal comparison serves as a point of departure for examining the issue of whether PROs are missing significant numbers of cases of inappropriate care, it confounds our two basic questions of appropriateness. To provide a more consistent basis for addressing the Subcommittee's second question, we first present information on rates of inappropriate use of specific diagnostic and surgical procedures detected in the PRO program and on those rates reported in the literature, followed by similar information on rates of inappropriate admissions.

Our data on the rates of inappropriate care in the Medicare program come from the four sources described in chapter 1:

- the published research literature,
- an independent review conducted for us of the appropriateness of the use of coronary angiography,
- data on PRO review activities tabulated by HCFA, and
- SuperPRO data.

Because SuperPRO, in theory, replicates the PRO review process, a direct comparison of the results of PRO and SuperPRO reviews is particularly useful in identifying potential reasons for observed differences in detected rates of inappropriate care. There are, however, some differences in the two processes that must be recognized. One is that SuperPRO nurse reviewers work off-site with copies of medical records rather than on-site in the hospital with the original records. Similarly, SuperPRO physician reviewers make their determinations without having any contact with the attending physicians. They may, however, receive any supplementary information provided to PROs by the attending physician, if PROs submit that information to SuperPRO. A second difference is that SuperPRO nurses are permitted to refer cases that they suspect of being inappropriate admissions to physician reviewers, even if the cases meet the written review criteria. While PRO nurses are also given this decision-making latitude, our belief is that they exercise this option less frequently than do SuperPRO nurses.

[1]See M. R. Chassin, et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?", Journal of the American Medical Association, 258(1987), pp. 2533-37.

The primary limitation that is shared by all but one of the individual studies or sources of information is that the results are not generalizable to the entire Medicare population.[2] We have already noted that PROs review a group of cases that is not representative of hospital care provided to Medicare beneficiaries. Since SuperPRO samples cases from each PRO's caseload, its group of cases is also not representative. In addition, the published studies on inappropriate admissions, although usually restricted to adults, are not always confined solely to Medicare beneficiaries. The published studies are also generally based on medical care provided prior to the implementation of the Medicare Prospective Payment System (PPS) and may not accurately reflect changes in hospital care brought about by PPS. As a result, it is possible to compare PRO denial rates to other estimates of inappropriate care in selected populations but not to provide a single estimate of inappropriate care in the Medicare program.

Despite these limitations, collectively, the information allows us to draw some conclusions about the performance of PROs in identifying inappropriate care in the Medicare program. The consistency of results across the various data sources increases our confidence that there are real differences between what PROs deny as inappropriate admissions and the actual extent of inappropriate care. The various sources of information cover a sufficiently diverse set of geographical areas and include enough post-PPS experience to make it unlikely that any single methodological weakness or potential "alternative explanation," including the implementation of PPS, explains any particular observed difference between PRO review results and those obtained by other reviewers.

[2]The Inspector General of the Department of Health and Human Services commissioned a study of a random sample of 7050 Medicare fiscal year 1985 hospital discharges. The study, conducted by the Health Data Institute, examined a broad range of issues, including DRG assignment, appropriateness of hospital admissions, and quality of care. See the Health Data Institute, National DRG Validation Study(Lexington, Mass.: The Health Data Institute, 1987).

Reviews of the Appropriateness of Individual Services

PRO Preadmission Reviews of Procedures

This section summarizes the results of PRO preadmission review of selected procedures, which were taken from a Project Hope survey of all PROs regarding their preadmission review activities.[3] In reviewing the data, it is important to note that a PRO prospective admission denial may involve judgments about both the need for the medical care and the appropriateness of the hospital setting. For example, for certain "inpatient" procedures (and DRGs related to those procedures), a denial of an inappropriate admission is tantamount to a judgment that the procedure was inappropriate. However, for procedures that can be done on an outpatient basis but sometimes require hospital admission, an admission denial cannot also be said to represent a judgment about procedure necessity. Instead, it may represent only a judgment that a hospital is not the proper setting within which to perform the procedure.

As shown in table 3.1, pacemaker insertions—the one procedure HCFA mandated for review under the second round of contracts—and cataract and lens procedures are the only procedures selected for preadmission review by more than half of PROs responding to the Project Hope survey. For each of the procedures listed, at least one PRO reported no admission denials. For several procedures, all PROs had less than 1 percent denials. Only potential outpatient procedures (that is, breast biopsy, bunionectomy, carpal tunnel release, cataract and lens procedures, dilatation and curettage, endoscopic procedures, and hernia repair) had mean admission denial rates in excess of 3 percent. This pattern suggests that the locus of treatment is currently more important in PROs' prospective determinations of appropriateness than the need for the procedure. That is, PROs concentrate more on where the treatment should be provided than on whether it is needed.

[3]See Project Hope, A Study of the Preadmission Review Process, Vol. 1(Chevy Chase, Md.: Project Hope, 1987).

Table 3.1: Hospital Admission Denial Rates for Procedures Frequently Designated for PRO Preadmission Review

Procedure	Number of PROs[a]	Distribution of denial rates			
		Minimum	Maximum	Mean	Standard deviation
Pacemaker insertion	40	0.0%	9.0%	0.7%	1.7%
Breast biopsy/procedures[b]	10	0.0	17.7	4.3	6.1
Bunionectomy[b]	4	0.0	18.6	9.8	7.8
Cardiac catheterization	6	0.0	6.4	1.6	2.8
Carotid endarterectomy	5	0.0	8.2	1.8	3.6
Carpal tunnel release[b]	15	0.0	31.0	11.5	10.3
Cataract and lens procedures[b]	24	0.0	25.7	3.2	5.6
Cholecystectomy	13	0.0	0.3	0.0	0.1
Coronary artery bypass graft	6	0.0	0.2	0.0	0.1
Dilatation and curettage[b]	13	0.0	16.7	7.1	6.4
Endoscopic procedures[b]	13	0.0	20.2	4.5	6.2
Hernia repair/procedures[b]	7	0.0	14.8	6.7	6.4
Hysterectomy	11	0.0	9.2	1.1	2.9
Hip/knee/ankle replacement	8	0.0	1.3	0.3	0.5
Transurethral prostatectomy	14	0.0	0.2	0.0	0.1

[a]A total of 40 PROs provided this information in response to a Project Hope survey for the Prospective Payment Assessment Commission.

[b]Potential outpatient surgical procedures based on HCFA ambulatory surgery list

Source: Project Hope, A Study of the Preadmission Review Process, Vol. 1 (Chevy Chase, Md.: Project Hope, 1987).

Published Studies

The published literature on the percentage of Medicare patients receiving inappropriate surgical procedures is limited to a few procedures and a few geographical areas. The studies that are available suggest that a sizeable proportion of some medical procedures may be inappropriate. Researchers affiliated with the Rand Corporation have published reports on the inappropriate use of carotid endarterectomy, coronary angiography, upper gastrointestinal tract endoscopy, and coronary artery bypass surgery. Researchers associated with the Philadelphia Professional Standards Review Organization have studied inappropriate pacemaker insertions. (See table 3.2.)

Table 3.2: Appropriateness of Care for Five Procedures

Procedure	Date of care	Sample size	Percent appropriate	Percent equivocal	Percent inappropriate
Coronary angiography[a]	1981	1,677	74%	9%	17%
Carotid endarterectomy[a]	1981	1,302	35	32	32
Upper gastrointestinal tract endoscopy[a]	1981	1,585	72	11	17
Coronary artery bypass surgery[b]	1979–82	386	56	30	14
Pacemaker insertion[c]	1983	382	44	36	20

[a]M. R. Chassin, et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?", Journal of the American Medical Association, 258 (1987), pp. 2533-37.

[b]C. M. Winslow, et al., "The Appropriateness of Performing Coronary Artery Bypass Surgery," Journal of the American Medical Association, 260 (1988), pp. 505-9.

[c]A. M. Greenspan, et al., "Incidence of Unwarranted Implantation of Permanent Cardiac Pacemakers in a Large Medical Population," New England Journal of Medicine, 318(1988), pp. 158-63.

In the Rand studies, a screening protocol was developed using the Rand approach described in appendix I. A random sample of medical records was abstracted for Medicare patients having carotid endarterectomy, coronary angiography, and upper gastrointestinal tract endoscopy in each of three geographical regions in 1981. The coronary artery bypass graft surgery study was done using a random sample of medical records of coronary artery bypass surgeries in a western state during 1979, 1980, and 1982; all of these cases came from three randomly selected hospitals and were not restricted to Medicare patients.

In the Philadelphia pacemaker study, a panel of physicians placed a range of diagnoses commonly used to justify pacemaker insertions into three appropriateness categories similar to the Rand categories. Diagnoses were assigned to categories based on the literature, standard cardiovascular practice, and the physicians' professional experience. The panel then reviewed a series of medical records of Medicare beneficiaries who received pacemaker insertions during the first six months of 1983 and placed the patients into the three appropriateness categories based on the information in the charts.

For each of the five procedures examined, a relatively large percentage of the cases were found to have been done inappropriately, ranging from 14 percent of coronary artery bypass surgery cases to 32 percent of carotid endarterectomies. In each instance, these estimates are likely to be somewhat conservative because an effort was made to give cases higher appropriateness ratings where the information was ambiguous.

The monetary costs associated with inappropriate procedures could be quite high. For example, at an average cost of $13,000 per carotid endarterectomy, the Rand researchers estimate that as much as $300 million is spent each year on inappropriate carotid endarterectomies alone. If all procedures had rates of inappropriate use similar to those found in the Rand work, and nothing changed either the rate or cost of specific procedures, the Rand researchers estimate the total annual cost of inappropriate procedures could run as high as $50 billion.

Our Replication of the Appropriateness of Coronary Angiography in the Medicare Population

To provide a more current estimate of the appropriateness of care and a replication of the Rand work for one procedure, we contracted with Value Health Sciences, Inc., to review the use of coronary angiography in a sample of Medicare cases, using an updated Rand-style protocol reflecting current medical practice in early 1987. The cases had all been previously reviewed by PROs and by SuperPRO. Of the 213 usable cases, 70 percent were found to be appropriate, about 10 percent were equivocal, and 20 percent were inappropriate.[4] This distribution is not significantly different from the distribution of appropriateness found in the 1981 Medicare sample. (See table 3.3.) The reasons for the inappropriateness determinations were also quite similar in the two samples. (See appendix II.)

Table 3.3: Appropriateness of Use of Coronary Angiography: a Comparison of Samples

Medicare case selection	No. of cases	Appropriate	Equivocal	Inappropriate
1981 sample[a]	1,677	74.0%	8.5%	17.4%
1987 sample[b]	213	70.0	9.9	20.2

[a]See M. R. Chassin, et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?", Journal of the American Medical Association, 258(1987), pp. 2533-37.

[b]Of PRO and SuperPRO cases, selected by us.

With two exceptions, the types of cases receiving coronary angiography in 1987 were very similar to those in 1981. For example, angina cases accounted for 46 percent of the instances where angiography was performed in 1981 and 44 percent of the cases in 1987. However, the use of coronary angiography during an acute myocardial infarction increased dramatically. In 1981, only 2 percent of the angiographies were performed during acute myocardial infarction; in 1987, 15 percent of the

[4]Of the 246 SuperPRO records we selected for review, 33 were excluded from analysis for the following reasons: (1) 9 did not have a coronary angiography, (2) 10 had an angiography performed as part of a percutaneous transluminal coronary angioplasty, and (3) 14 did not have sufficient clinical data to establish a precise indication.

procedures occurred in this situation. The second change involved a drop in the proportion of patients being evaluated for valvular disease, from 19 percent in 1981 to 10 percent in 1987.[5]

Discussion

The available data on the extent of inappropriate use of specific procedures in the Medicare population are very limited. However, they suggest that a sizeable proportion of procedures may be done for inappropriate reasons, at a considerable cost to the Medicare program. These cases were not routinely labeled as inappropriate by PROs under the second round of contracts. However, the new emphasis, under the third round of contracts, on reviewing the need for invasive procedures in PRO retrospective reviews could result in an increase in denials of inappropriate procedures.

Reviews of the Appropriateness of Hospital Admissions

HCFA Data

Nationally, PRO denial rates for inappropriate admissions, both retrospective and preadmission, are very low. Over roughly a two-year period, 1986 through 1988, only 2.1 percent of admissions were denied retrospectively. While there is some variation across PROs in denial rates, individual PRO retrospective denial rates are all below 6 percent, and only 9 states have denial rates at or above 3 percent. (See table 3.4.)

[5]PROs denied admission in 2 percent of the coronary angiographies found inappropriate in the review conducted for us and in no case in either of the other two categories. SuperPRO reviews questioned the appropriateness of admission in a greater percentage of cases; according to our results, the percent of admissions determined to be inappropriate increased as the use of the procedure became more questionable. Neither PROs nor SuperPRO focused as directly on the issue of the appropriateness of the coronary angiography as did Value Health Sciences.

Table 3.4: Percent of Reviewed Hospital Admissions Denied by PRO

PRO	Admission denials	
	Preadmission[a]	Retrospective[b]
Alabama	0.3%	2.0%
Alaska	3.0	2.7
Arizona	0.4	2.0
Arkansas	0.2	2.1
California	2.2	2.5
Colorado	0.5	2.0
Connecticut	0.2	3.2
Delaware	0.0	2.0
District of Columbia	4.2	4.2
Florida	0.2	1.8
Georgia	0.2	2.0
Guam/American Samoa	[c]	1.5
Hawaii	0.0	0.8
Idaho	3.7	1.6
Illinois	1.9	4.4
Indiana	0.9	2.4
Iowa	0.7	1.5
Kansas	0.3	3.5
Kentucky	4.3	2.8
Louisiana	2.8	1.7
Maine	0.0	2.3
Maryland	8.0	1.8
Massachusetts	0.5	2.0
Michigan	1.8	2.6
Minnesota	0.4	1.5
Mississippi	0.1	2.6
Missouri	0.0	3.0
Montana	0.8	2.0
Nebraska	1.0	2.9
Nevada	0.0	3.7
New York	3.4	1.3
New Jersey	0.0	1.2
New Mexico	0.3	2.9
New Hampshire	0.2	2.1
North Carolina	0.0	1.4
North Dakota	0.0	2.5
Ohio	0.0	1.7
Oklahoma	0.4	0.6
Oregon	8.5	2.4

(continued)

PRO	Admission denials	
	Preadmission[a]	Retrospective[b]
Pennsylvania	1.2	1.8
Puerto Rico	0.0	1.2
Rhode Island	1.6	1.8
South Carolina	0.0	5.7
South Dakota	0.2	1.3
Tennessee	0.8	2.2
Texas	0.0	2.4
Utah	0.7	3.2
Vermont	0.3	1.2
Virgin Islands	[c]	4.3
Virginia	0.2	1.7
Washington	0.4	1.2
West Virginia	0.0	2.9
Wisconsin	0.0	1.8
Wyoming	0.0	1.5
National aggregate	0.6%	2.1%

[a]Denial rates for preadmission reviews are calculated from the number of cases reviewed and the number of denials for the data elements representing pacemakers (p. B-5), DRGs/procedures (p. B-6), and "other" preadmission reviews (p. B-7) in the HCFA/Health Standards and Quality Bureau/Office of Peer Review PRO data summary.

[b]Retrospective rates refer to the percent of unduplicated, reviewed cases for which admission was denied for utilization (as opposed to coverage) reasons.

[c]Preadmission reviews were not conducted by PRO in America Samoa/Guam and the Virgin Islands.

Source: HCFA/Health Standards and Quality Bureau/Office of Peer Review; PRO data summary reflecting data accepted through February 1989. Data collected from monthly PRO reports to HCFA under the Second Scope of Work, roughly 1986-1988.

Only 0.6 percent of cases reviewed prospectively were denied.[6] However, the variation in denial rates for preadmission reviews is considerably greater than that for retrospective review. Fifteen states report no preadmission denials, while Maryland and Oregon denied 8.0 and 8.5 percent of admissions respectively. An examination of the results of Project Hope's survey, discussed previously, indicated that states with higher preadmission denial rates focus their preadmission reviews on potential outpatient surgical procedures or specific types of medical admissions, such as back problems.

[6]These data reflect reviews done under the second round of contracts. During that time, PRO preadmission review was limited to pacemaker implants (all PROs) and four other procedures (selected independently by each PRO).

Superpro Data

Overall Findings

When SuperPRO replicated the PRO retrospective reviews of admission necessity on a random sample of cases, they found a considerably higher percentage of inappropriate admissions than PROs did. In the case of every PRO except Nevada and West Virginia, SuperPRO would have denied care for at least twice as many admissions as PRO did. (See table 3.5.)

Table 3.5: Comparison of Retrospective Admission Denials for Cases Reviewed by Both PRO and SuperPRO, by PRO

PRO	Number of cases	Denied by PRO[a]	Denied by SuperPRO
Alabama	1,144	2.5%	10.5%
Alaska	934	1.6	16.5
Arizona	1,140	1.7	10.2
Arkansas	1,160	4.6	10.9
California	1,192	1.1	6.8
Colorado	1,136	2.6	9.0
Connecticut	1,148	3.6	11.0
Delaware	942	2.8	12.6
District of Columbia	844	4.0	17.3
Florida	973	3.5	17.7
Georgia	1,135	3.0	13.0
Guam/Samoa	36	.0	36.1
Hawaii	684	.9	9.6
Idaho	668	.4	8.1
Illinois	1,176	3.5	15.8
Indiana	1,164	2.4	8.7
Iowa	1,155	1.4	10.1
Kansas	1,089	3.3	13.5
Kentucky	1,142	2.5	10.9
Louisiana	1,147	2.6	13.1
Maine	1,092	1.8	11.2
Maryland	772	1.9	14.8
Massachusetts	721	3.5	11.4
Michigan	778	3.5	15.2
Minnesota	1,067	.9	10.9
Mississippi	1,158	2.8	16.7
Missouri	1,173	3.1	12.8
Montana	1,015	.7	17.8
Nebraska	1,087	5.1	15.1
Nevada	997	8.3	13.4

(continued)

PRO	Number of cases	Denied by PRO[a]	Denied by SuperPRO
New York	1,101	1.9	10.1
New Mexico	945	1.2	12.7
New Jersey	1,166	1.4	9.3
New Hampshire	1,012	1.5	13.6
North Carolina	1,170	1.8	9.3
North Dakota	1,073	1.4	6.6
Ohio	790	1.5	12.4
Oklahoma	1,161	1.6	9.7
Oregon	1,128	3.5	10.7
Pennsylvania	785	3.3	11.7
Puerto Rico	729	1.9	14.5
Rhode Island	1,045	1.6	10.5
South Carolina	765	5.2	16.5
South Dakota	1,111	2.0	17.0
Tennessee	1,158	1.7	12.1
Texas	1,178	4.6	11.4
Utah	1,079	1.7	8.4
Vermont	1,045	1.1	12.2
Virgin Islands	159	8.8	28.3
Virginia	1,167	2.4	10.6
Washington	1,161	1.3	9.1
West Virginia	1,147	5.1	9.8
Wisconsin	1,129	3.7	14.4
Wyoming	869	.7	17.3
National aggregate	53,942	2.6%	12.1%

[a]PRO-denial percentages reported here differ from those in table 3.4 because this table includes only those PRO cases subsequently rereviewed by SuperPRO.

Source: Our analysis of SuperPRO data

Potential Explanations for Differences

In attempting to understand why SuperPRO finds a considerably higher rate of inappropriate admissions than do PROs, we explored a number of potential explanations, including differences in nurse referral patterns and differences in final physician judgments.

Nurse Referrals. Because SuperPRO nurses are, at least nominally, using the same screening protocols for admission necessity as PRO nurses, agreement on referrals to physicians for review should be relatively high. If referral rates differ and if poor reliability in applying the screening criteria is the only factor involved, any disagreements on referrals should be equally split between cases where only the PRO nurse

recommends referral and those where only the SuperPRO nurse recommends referral. Because PRO and SuperPRO physicians generally review admission necessity only on those cases that are referred to them, systematic differences between PRO and SuperPRO nurses in cases screened out for review could be one reason for the observed differences in final physician determinations.

Overall, the nurses agreed on 81.4 percent of the cases that both reviewed—73.1 percent of the cases met the admission criteria, and 8.3 percent failed and thus were referred for physician review. (See table 3.6.) This 81.4 percent figure represents a reasonably high rate of agreement. However, of the remaining 18.6 percent of cases on which the nurses disagreed, SuperPRO nurses referred over five times as many cases for failing to meet the PRO admission criteria as did the PRO nurses—that is, SuperPro nurses referred 15.6 percent versus 3.0 percent for PRO nurses. This suggests that SuperPRO nurses apply the PRO protocols more stringently than do PRO nurses.

Table 3.6: PRO and SuperPRO Nurse Referrals for Physician Admission Necessity Review

	SuperPRO nurse			
	Referred		Did not refer	
PRO nurse	Number	Percent	Number	Percent
Referred	4,272	8.3%	1,526	3.0%
Did not refer	8,012	15.6	37,534	73.1

For the 3553 cases referred by both PRO and SuperPRO nurses for admission necessity review for which a SuperPRO physician judgment was recorded, the SuperPRO physician agreed in 72.4 percent of the cases that the admission was inappropriate.[7] For the 84 cases referred only by PRO nurses, 66.7 percent were judged inappropriate. For the 5184 cases referred only by SuperPRO nurses, 54.6 percent were judged to be inappropriate. For the 598 cases on which both the PRO and SuperPRO nurses agreed that referral for admission necessity review was not required, but which were referred to a SuperPRO physician for other reasons, the SuperPRO physician judged admission to be unnecessary in 3.2 percent of the cases. Although this percentage is low, it is higher than the overall percentage of admissions found inappropriate in the original PRO review—2.1 percent.

[7]The numbers of cases in this section differ from those in the last section because a number of cases did not have final SuperPRO physician determinations recorded in the data base. For cases that are not referred by the SuperPRO nurse for any reason, a final SuperPRO physician judgment would not be expected. However, in other cases where a judgment would be expected, we could not determine the reason for the absence of final SuperPRO physician determinations.

These results show that while the percent of cases found to be inappropriate admissions is somewhat lower for cases referred only by SuperPRO nurses than for cases referred only by PRO nurses, over half of the extra cases identified by the SuperPRO nurses were still determined to be inappropriate admissions. Further, SuperPRO nurses were more likely to refer cases than were PRO nurses. In addition, a percentage of inappropriate admissions at least as large as that denied by PROs is not being identified by the screening protocols, even as applied by the SuperPRO nurses.

"Looseness" of the PRO Screening Protocol. One indicator that the PRO criteria as applied by the nurse reviewers miss some inappropriate admissions comes from the SuperPRO physicians' determination that sizeable proportions of cases are inappropriate, even though either the PRO or SuperPRO nurse felt that the screening criteria were met. A second test of the potential looseness of the criteria comes from SuperPRO physician judgments on a group of cases that SuperPRO nurses referred because they questioned the appropriateness of hospital admission, despite the fact that the cases formally met one of the PRO's admission criteria. A total of 2,598 cases were referred to SuperPRO physicians on this basis. (These cases were not included in prior analyses.) Of those 2,265 cases for which a SuperPRO physician judgment is recorded, 47.4 percent were judged to be inappropriate admissions.

Physician Judgments of Appropriateness. Of the 4,053 cases on which both a PRO physician and a SuperPRO physician made a judgment, they agreed on 56.9 percent of the cases (30.2 percent of the admissions judged inappropriate; 26.7 percent judged appropriate). Virtually all of the disagreements occurred in cases where the SuperPRO physician judged a case to be an inappropriate admission and the PRO physician judged it to be appropriate. (SuperPRO physicians found 41.2 percent of the admissions to be inappropriate; PRO physicians found only 1.8 percent of them to be inappropriate.)

Published Data

Overall Findings

The published literature on the extent of inappropriate hospital admissions in general has been reviewed in two recent articles.[8] Table 3.7 presents only those studies that include at least some Medicare patients, although younger adults may have been studied as well. Where possible, separate estimates for the Medicare population are presented.

[8]See S.M.C. Payne, "Identifying and Managing Inappropriate Hospital Utilization: A Policy Synthesis," Health Services Research, 22:5(1987), pp. 710-69, and R. H. Brook, et al., "Appropriateness of Acute Medical Care for the Elderly," Rand manuscript in press, 1989.

Table 3.7: Studies of Inappropriate Hospital Admissions

Study[a]	Study location	Dates of care	Study population	Review protocol[b]	Percent of inappropriate admissions
Restuccia, et al., 1984	4 PSROs	1980	1,232 adult admissions	AEP	19%[c]
Borchardt, 1981	Delmarva PSRO	1980–81	2,711 Medicare and Medicaid admissions	AEP	14
HCFA, 1984	National	1981	5,732 adult admissions	SMI	6; 4.3[d]
Restuccia, et al., 1987	1 hospital, Massachusetts	1981–82	297 adult admissions	AEP	12
Health Data Institute, 1984	Delmarva PSRO	1982	2,085 Medicare admissions	AEP	10
Studnicki and Stevens, 1984	6 hospitals, Maryland	1983	467 Medicare discharge	AEP	7
Restuccia, et al., 1986	7 New England hospitals	1983–84	12,071 adult admissions	AEP	12
Health Data Institute, 1987[e]	National	1984–85	7,050 Medicare discharges	AEP	10

[a]J. D. Restuccia, et al., "A Comparative Analysis of the Appropriateness of Hospital Use," Health Affairs (Summer 1984), pp. 130-38.

P. J. Borchardt, "Nonacute Profiles: Evaluations of Physicians' Nonacute Utilization of Hospital Resources," Quality Review Bulletin (November 1981), pp. 21-26.

Health Care Financing Administration, National Estimate of Nonacute Hospital Utilization for 1981 (Washington, D.C.: U.S. Department of Health and Human Services, 1984).

J. D. Restuccia, et al., "Assessing the Appropriateness of Hospital Utilization to Improve Efficiency and Competitive Position," Health Care Management Review (Summer 1987), pp. 17-27.

Health Data Institute, Analysis of DRG Patterns and Appropriateness of Utilization in Delmarva (Newton, Mass.: Health Data Institute, 1984).

J. Studnicki and C. E. Stevens, "The Impact of a Cybernetic Control System on Inappropriate Admissions," Quality Review Bulletin (October 1984), pp. 304-11.

J. D. Restuccia, et al., Reducing Inappropriate Use of Inpatient Medical/Surgical and Pediatric Services (U.S. Department of Commerce: National Technical Information Services, 1986).

Health Data Institute, National DRG Validation Study (Lexington, Mass.: Health Data Institute, 1987).

[b]AEP = Appropriateness Evaluation Protocol
SMI = Standardized Medreview Instrument

[c]The range was 12 to 28 percent across the four Professional Standards Review Organizations (PSROs).

[d]6 percent for all payers; 4.3 percent for medicare patients

[e]This is the study referred to as the Department of Health and Human Services Inspector General's study elsewhere in the report. The estimate of inappropriate admissions is the weighted national estimate.

With one exception, all of the studies used the Appropriateness Evaluation Protocol (AEP) to determine the appropriateness of admissions. The estimates of inappropriate admissions in those studies range from 7 to 28 percent. However, estimates for the Medicare population alone are generally lower, ranging from 7 to 10 percent.

The only study on inappropriate admissions in the post-PPS environment, the National DRG Validation Study, was conducted for the Department of Health and Human Services Inspector General by the Health Data Institute. It is also the only study in the group with a nationally-representative sample of cases. In that study, 10 percent of Medicare discharges were found to have been inappropriate admissions because they did not meet any of the AEP criteria for admission. Based on the DRG assignment of the specific cases found inappropriate and the related hospital-specific payment rates, the study estimated that inappropriate admissions cost the Medicare program about $2 billion, or 7.4 percent of the $27 billion spent on PPS in fiscal year 1985.

The study also found that an additional 34.4 percent of cases met only one AEP criterion for admission. In 57.5 percent of these cases, the only AEP criterion met was the need for intravenous therapy. The study notes that while some intravenous therapy requires supervision in an acute care setting, much of it can be safely provided in the home or other outpatient setting. The only other individual criteria accounting for more than 8 percent of the cases were vital sign monitoring (8.6 percent) and entry into the body cavity (8.5 percent).

Potential Explanations for Differences

The rate of inappropriate hospital admissions based on PRO reviews (that is, 2.1 percent) is generally lower than the rate of inappropriate admissions reported in the published studies (that is, approximately 10 to 12 percent). The time period in which the studies were done does not appear to be a viable explanation because the only national study done in the post-PPS environment found roughly the same level of inappropriate admissions as did the earlier studies. One difference that should be considered is the screening protocol used. In almost all of the published research, AEP is used; most PROs use the ISD-A system or some modification of it. However, the possibility that different screening protocols account for the differential rates of inappropriate admissions is lessened by similar discrepancies between the results of PRO and SuperPRO reviews. In the latter instance, the same review protocols are used. More specifically, our analysis of SuperPRO data suggests that the differential application of the review protocol is probably more important than the specific protocol used.

The only other apparent difference is in the context and implications of the final determination that an admission is inappropriate. In the published literature, as in the SuperPRO review, the final determination is based primarily on a review of the written record—without any requirement to interact with the attending physician—and no monetary

penalties result from the review. By contrast, PROs contact the attending physician when questions of appropriateness arise and thus are required to interact directly with local practitioners. PRO reviewers are themselves part of that professional community and must maintain their standing within it. Undoubtedly, the interaction with the attending physician concerning questionable care can reveal extenuating circumstances that justify the initial plan of treatment. In such cases, the additional information may well serve the patient's interest. In other cases, the interaction may facilitate the creation of a rationale for not denying payment even though the admission remains truly questionable. Regardless, the roles and relationships of PRO reviewers are more complex than those of either SuperPRO reviewers or researchers. This may affect their review findings.

Summary and Conclusions

PROs have only recently begun to systematically review the appropriateness of individual procedures retrospectively. While they have prospectively reviewed the appropriateness of procedures since the beginning of the second round of contracts, the reviews have been limited to a small number of primarily inpatient procedures. In both instances, PROs have rarely questioned the appropriateness of care. The published literature, supplemented by our own study of the inappropriate use of coronary angiography, finds higher rates of inappropriate care. This suggests that additional PRO attention to reviewing the appropriateness of individual surgical procedures is warranted.

Our review of various estimates of the rate of inappropriate admissions in the Medicare program shows that PROs, in comparison with either SuperPRO or published research results, consistently discover a lower rate. Some potential explanations for these differences seem more likely than others. The judgments of admission necessity in the published literature are frequently based on the AEP review protocol, but many PROs use the ISD-A system as the basis for screening cases. Thus, the screening protocol itself could be a potential explanation. However, given that SuperPRO uses the same protocols as PRO and still finds more inappropriate care, the screening protocol is not likely to be a major factor in the differences in inappropriate admission rates. A differential application and interpretation of the protocols could be a significant issue.

One major difference between the published literature and PRO reviews of inappropriate admissions is the types of cases reviewed. While much of the literature does not use nationally-representative samples of Medicare cases, the samples are randomly drawn from Medicare discharges

in the hospitals in the limited geographical regions studied. As we indicated in chapter 2, the types of cases reviewed retrospectively by PROs are, with the exception of the 3-percent random sample, not random samples of hospital discharges in their areas. Although the admission denial rate in PROs' 3-percent random sample is higher than that for many of the other categories of cases reviewed, it is still considerably lower than that found by researchers. In addition, this explanation does not account for the considerably higher rates of inappropriate admissions found by SuperPRO, since SuperPRO randomly samples cases reviewed by PROs.

One possible reason for fewer PRO admission denials that applies to both the published literature and the SuperPRO reviews is that PROs, and their physician reviewers, feel strongly that they need to maintain a positive relationship with the local medical community. Philosophically, some PROs believe that the best way to change physician practice patterns is by providing feedback on current practice patterns and initiating educational interventions, rather than by denying payment. For the PRO system to be effective, they believe an interactive, collegial working relationship is required. In addition, at an individual level, PRO physicians are in the position of trying to enforce the Medicare requirements, which have financial and professional implications for the provider under review, while at the same time maintaining their own personal and professional relationships in their communities. Thus, PRO reviewers may have less incentive to confront their colleagues about providing inappropriate care than would reviewers whose judgments do not have immediate reimbursement implications or who are more removed from the local medical community.

Approaches to Strengthening Utilization Review in the Medicare Program

Introduction

The Subcommittee asked us not only to provide information describing approaches to UR and the level of inappropriate hospital care in the Medicare program but also to use that knowledge to suggest ways of improving Medicare UR. In this chapter, we draw on the information presented in chapters 2 and 3 to make such suggestions.

Our suggestions for improvement are driven by the finding that PROs are discovering a lower rate of inappropriate hospital admissions for Medicare beneficiaries than are any of the other review or research activities we examined. The more limited information on the appropriateness of individual procedures suggests a similar conclusion. Inappropriate use of hospital services means that the beneficiary may be unnecessarily subjected to the risks associated with hospitalization and with potentially unnecessary diagnostic and therapeutic procedures. Furthermore, the cost to the Medicare program of the as-yet-unidentified inappropriate care could be millions, if not billions, of dollars annually. Thus, improvements in PRO UR activities could be advantageous to both the beneficiary and the Medicare program.

Identifying the reasons for the differences in rates of inappropriate care and using that information as the basis for recommending improvements is not a straightforward task. Except in a small percentage of cases where the evidence is undeniable, it seems to be very difficult for PROs to aggressively question the appropriateness of hospital care. This tendency may be inherent in professional peer review. Perhaps if the review of care were removed from the local medical community and performed by review physicians accountable to a more autonomous UR organization, it would be easier to raise questions of appropriateness. However, the principles underlying the PRO program assume the support and endorsement of the local medical community. Beyond this, characteristics of the PRO review process itself may reduce the likelihood that inappropriate care will be identified.

Private Sector UR Activities May Offer Some Useful Alternatives to the PRO Approach

The primary difference between private sector and Medicare UR is the timing of the review. Three related aspects of private sector prospective reviews also seem important: (1) targeting reviews toward cases with increased probability of inappropriate care, (2) the advisory nature of prospective review decisions, and (3) structuring health benefit plans to encourage compliance with UR programs.

Prospective Review

Most PRO reviews of the appropriateness of medical care take place after the care has been rendered—that is, retrospectively. When a hospitalization is determined to be inappropriate, it means that the beneficiary has unnecessarily been subjected to the risks of hospitalization, as well as the risks of unnecessary diagnostic and therapeutic procedures. Furthermore, the automatic decision to deny payment has important financial implications, since resources have already been committed to care for the patient. In contrast, prospective review provides an opportunity to assure that Medicare beneficiaries receive the care they need without being subjected to the medical and financial risks of inappropriate hospital care.

The available information on the cost-effectiveness of prospective review, although limited, clearly indicates that the private sector's experience is positive. Because the Medicare population is older, however, and probably has more comorbidities (that is, multiple health problems occurring simultaneously), Medicare patients may be more likely to require hospitalization than do younger, healthier patients. If this is the case, the cost-effectiveness of prospective review in the Medicare program could be less than it is in the private sector. Thus, the cost-effectiveness of various options for expanding prospective reviews would require careful examination before additional prospective review requirements were instituted. However, given the sizeable percentage of questionable Medicare admissions, it is our opinion that, with appropriate targeting, prospective review could be cost-effective in Medicare as well.

With regard to what aspects of appropriateness to review, both the appropriateness of particular procedures and the appropriateness of the hospital setting should be considered. Implementing the second-opinion program already required by the Congress could be a useful adjunct to

prospective review of procedure necessity.[1] Because of the incentives of PPS, length of stay review appears less useful.

Better Targeting of Reviews

The data presented in chapter 2 make it clear that some of the categories of cases frequently determined to be inappropriate admissions are not currently targeted in the PRO retrospective review process. For example, elective admissions for a surgical procedure generally performed in an outpatient setting would rarely fall into one of PRO's mandated categories of cases. In addition, certain medical admissions that are prone to be inappropriate would rarely be reviewed. While the types of cases reviewed retrospectively by PROs reasonably reflect the variety of review tasks assigned to them, most of the cases reviewed fall into categories with relatively low rates of inappropriate care. The data presented in chapter 2 demonstrate that higher rates of denial are obtained when hospitals, fiscal intermediaries, and PROs specifically target cases for appropriateness review. The research literature also provides additional useful information for targeting reviews of appropriateness. To increase the effectiveness of UR, cases that frequently have been shown to be inappropriate admissions might be targeted for prospective review. Further, it may make sense to review all elective admissions, as is done in the private sector. If targeting reviews toward specific providers who are delivering inappropriate or unnecessary care is a high priority, a representative sample of each hospital's discharges would have greater utility in identifying providers for focused investigations than the particular mix of cases currently reviewed by PROs.

Advisory Nature of Prospective Determinations

The prospective UR decision in the private sector is typically advisory. The patient and physician are advised that the proposed care may not be paid for if retrospective review indicates that it was inappropriate. As a result, the input of the review physician becomes one more piece of information that the patient and the attending physician have to consider in making a decision on the final treatment plan.

Further, our interviews with individuals involved in private sector UR programs suggest that the threat of noncertification for payment is only one of the deterrents associated with prospective private sector UR.

[1] Although the Consolidated Omnibus Reconciliation Act of 1985 requires second opinions in conjunction with the current prospective review of ten surgical procedures, HCFA has not yet implemented this requirement. Second opinions would be required when the medical appropriateness of any of the procedures was questioned.

They emphasize the sentinel, or deterrent, effect of having to obtain certification and the use of persuasion by review physicians to encourage attending physicians to use alternative treatments and settings. In addition, they note that because the review is usually advisory—providing an opportunity for discussion and persuasion—reviewers may find it easier to question proposed treatments than they might in the Medicare program, where intervention options are more limited and a decision that care is inappropriate leads to automatic denial of payment. We believe that HCFA should consider these types of private sector activities—perhaps initially on a demonstration basis. If HCFA should subsequently propose to make some prospective UR decisions advisory, legislative changes would almost certainly be required.

Tools to Gain Compliance With Prospective Determinations

Many health benefit plans in the private sector have been structured to provide a penalty for not complying with the requirement to obtain preadmission certification or an incentive for complying with the UR advice. For example, copayment and deductible amounts may be altered depending on whether the treatment rendered is consistent with a prospective UR judgment. This is also an area where experimentation could be useful, although legislative changes would be required to make such provisions a permanent feature of the Medicare program.

Additional Options to Consider

Given the extensive use of prospective review in the private sector, contracting with a limited number of firms that specialize in prospective utilization review might represent a cost-efficient approach for the Medicare program. Such an approach might also produce more uniform results than would be obtained if 54 different PROs did the reviews, particularly if a more standardized set of review protocols were adopted as well.

With regard to protocols, both PROs and the private sector use explicit protocols such as AEP and ISD-A, as well as locally-developed diagnostic-specific criteria, to screen cases manually for appropriateness review. One option to consider is greater reliance on national standards for appropriate care—at least in those areas where substantial agreement exists on preferred treatment options. If national standards are found desirable, HCFA may lack the necessary statutory authority to use them in peer review. Although one may argue that the use of such protocols for screening cases, with final determinations based on physician reviews, would satisfy the requirement for taking local practice variations into account, clear statutory authority to adopt national standards

may be desirable. If so, HCFA should request that authority from the Congress.

In the event that the protocols are prospectively applied to a large volume of cases in the future, the feasibility of computerizing them in order to streamline the review process might then be explored. Many private sector utilization review companies have automated their prospective reviews of admission necessity and their length-of-stay determinations. Some companies that also examine the need for specific procedures have automated that screening activity as well. A greater reliance on automated screening in prospective review would necessitate more standardization in the protocols used by PROs than has been the case to date.

Improvements in Current PRO Retrospective Reviews Are Feasible, With or Without Expanded Prospective Review

If prospective review were expanded in the PRO program, current retrospective review requirements would need to be reexamined to avoid redundancy and to increase the efficiency and effectiveness of both types of review. Shifting the review of appropriateness in individual cases to a prospective mode suggests the need to reconsider how retrospective reviews of appropriateness should be handled. For example, some retrospective medical record reviews would be needed to check on the accuracy of information provided to the review organization during the prospective review process. Also, some retrospective manual review of medical records would continue to be necessary to fulfill other PRO review requirements, such as confirming problems with quality of care and checking on the accuracy of DRG assignments. However, it would not make sense to do retrospective appropriateness reviews on all of the currently reviewed cases.

Regardless of the extent of expansion of prospective review, the use of more standardized protocols for screening care, the automation of some of the initial screening of care, and better targeting of appropriateness reviews could lead to more efficient use of review resources and a fuller identification of inappropriate care. The successful development of HCFA's Uniform Clinical Data Set would also enhance the ability of PROs to efficiently conduct their review activities in these areas.

Better targeting of retrospective appropriateness reviews, in conjunction with the expansion of prospective reviews, suggests that increased reliance on profiling might be effectively used to focus or target retrospective reviews on categories of patients that exhibit significant variation in the care provided or on providers whose patterns of care are dramatically different from those of their peers. The type of profiling

required to accomplish this sort of targeting would involve looking at the patterns of care associated with particular types of cases (for example, diagnoses or surgical procedures) or individual providers over time. More detailed information than is currently collected would be required to develop the profiles based on patterns of care and to ensure that differences between providers in regard to the types of patients they treat are not misidentified as utilization problems. The Uniform Clinical Data Set, which is currently being developed and field-tested by HCFA, represents one approach to obtaining such additional information.

Once the profiling activities identified specific types of patients or individual providers as targets for review, only the medical records of those patients or providers would then need to be subjected to a more intensive, case-by-case review in order to confirm whether a truly anomalous practice pattern existed. If the focused reviews indicated that a problem did exist, PRO physicians would be in a better position to intervene directly through educational and other methods to improve the practice patterns of those providers. Those providers might also become the objects of more intensified prospective review.

In the short run, if all of the options suggested in this chapter were implemented, they could require more resources and expertise than many PROs currently possess. One approach would be to provide PROs with the resources necessary to obtain the needed hardware and expertise. Alternatively, some of the activities, particularly those related to data sources, could be maintained and analyzed centrally by HCFA, or by a contractor, either of whom could forward the results to PROs for input into local review activities. This might be another area for demonstration projects and would be worth pursuing as a means of improving retrospective review, even in the absence of an expansion of PRO prospective review activities.

In summary, we believe that a number of approaches to UR might be effective in reducing the level of inappropriate care in the Medicare program. While most of these approaches revolve around an expansion of prospective review activities, others involve changes in the current retrospective review requirements. Some of the changes could be implemented quite readily; others would require an examination of the cost-effectiveness of various options; and a few may require legislative changes.

Recommendation

We recommend that the Secretary of Health and Human Services require the Administrator of HCFA to expand prospective utilization review in the Medicare program. This step should be preceded by a systematic examination of the cost-effectiveness of various options, including those described in this report. If HCFA lacks legislative authority to implement particular options—such as advisory prospective determinations and the use of incentives or penalties for complying with prospective review decisions—it should seek that authority from the Congress. It is also important to evaluate the current retrospective review requirements in light of the expanded prospective reviews, in order to avoid unnecessary duplication and improve the efficiency and effectiveness of both types of review.

Agency Comments and Our Response

Department officials generally agreed with our recommendation to expand prospective review in the Medicare program. However, in their view, they have already expanded prospective review in the third PRO Scope of Work and intend to examine the results of these reviews and change the program as needed based on these results. Further, they do not believe that advisory opinions would be appropriate for Medicare because they feel that an advisory opinion would be less effective in forcing attending physicians to change their behavior than outright denials. They noted that some of the options we suggest would not be possible under the current statute. They also provided a number of technical comments. (See appendix IV.)

As we have described in this report, HCFA did expand prospective review from five to ten procedures in the third PRO Scope of Work. We view this expansion as minimal. Further, past review results suggest that very little care will be found to be inappropriate, due to the HCFA focus on surgical procedures done predominantly in an inpatient setting. We believe that increased emphasis on surgical cases normally treated in an outpatient setting and on medical cases for which hospitalization typically is not warranted would increase the cost-effectiveness of prospective review. Thus, we are not convinced that an evaluation of the prospective review requirements in HCFA's third PRO Scope of Work will adequately demonstrate the potential cost-effectiveness of prospective review. Finally, an evaluation of the current approach to prospective review will not enable the Department of Health and Human Services to adequately assess the cost-effectiveness of the various promising options discussed in this chapter.

We do not agree that advisory determinations are inappropriate in Medicare. If, as the Department suggests, one views this particular option as a way of forcing attending physicians to use alternative modes of treatment, then an advisory determination could be less useful than a payment denial. However, if one accepts our position that PRO physicians need "incentives" to more actively question the care provided by attending physicians than they have in the past, the advisory determination becomes a promising option. Finally, we agree, and make clear in the report, that some options we discuss would require legislative changes. (See pages 57 to 58.)

Availability of Protocols in the Private Sector

Diagnosis-Specific Protocols

Diagnosis- or procedure-specific protocols focus on the medical services considered appropriate for particular diagnoses, patient conditions, or procedures. For example, protocols might be developed for reviewing diabetes (a diagnosis), chest pain (a "condition"), or carotid endarterectomy (a procedure). In each case, the protocols would indicate which medical services are appropriate for patients in each of those categories.

Examples

Examples of diagnosis-specific protocols include those developed in the 1970's by the American Medical Association. More recently, the Rand Corporation used a more systematic methodology to develop six protocols (hereafter referred to as the "Rand protocols") as part of their larger Health Services Utilization Study. In addition, other protocols have been developed through the American College of Physicians' Clinical Efficacy Assessment Project and by the Blue Cross and Blue Shield Association's Medical Necessity Program.

The American Medical Association protocols were developed primarily for use in the Professional Standards Review Organization program, the precursor to the current PRO program for Medicare. The protocols were lengthy and listed the diagnostic tests and therapeutic services that were consistent with a given diagnosis. They did not attempt to differentiate between those that were appropriate for a given patient and those that were not. Many experts believe that these protocols were overly inclusive and not particularly useful for making judgments about appropriateness in individual cases. However, it is also the case that the development of protocols is methodologically difficult and that the American Medical Association protocols represent a pioneering effort.

In contrast, Rand used a highly structured approach aimed at specifying the instances in which particular procedures are clinically appropriate. First, on the basis of a review of the medical literature, the clinical indications for which a specific procedure had been used were identified and empirical evidence of the procedure's risks and benefits for that particular indication evaluated by Rand researchers. On the basis of this literature review and their own clinical judgment, a panel of expert physicians then rated each indication on a nine-point scale of appropriateness. The question of whether a procedure was appropriate was treated in terms of patient risks versus patient benefits of using the procedure, rather than in terms of procedure costs versus patient risks and benefits. Indications were labeled appropriate, equivocal, or inappropriate based on the median appropriateness rating and degree of agreement among the judges.

In evaluating the appropriateness of care for individual patients, using the Rand protocols, medical records are reviewed for all potential indications that might apply to that case. The case is placed into one of the three categories based on the indication substantiated by the medical record that is closest to "appropriate." For example, if the medical record had evidence to support an "equivocal" and an "appropriate" indication, the case was placed in the "appropriate" category.

The purpose of the American College of Physicians' Clinical Efficacy Assessment Project is to determine the medical merit of selected practices in internal medicine and disseminate that information to practicing physicians.[1] Judgments about medical merit are based on the available evidence concerning the safety and efficacy of the medical service. The protocols are typically developed by individual physicians (or small groups of physicians), reviewed and adopted by the executive committee of the American College of Physicians' board, and published in the Annals of Internal Medicine. The published protocols typically specify the circumstances in which a test or procedure is indicated, not indicated, or contraindicated.

Beginning in 1977, the Blue Cross and Blue Shield Association's Medical Necessity Program, in cooperation with national medical organizations, developed a number of protocols for assessing the medical necessity of particular diagnostic and therapeutic procedures. Currently, the protocols are based on reviews of the published scientific literature conducted by recognized medical experts and frequently supplemented by input from meetings with representatives of the relevant medical specialty societies. The program's first product was a list of outmoded or replaced procedures for which Blue Cross and Blue Shield plans were advised not to reimburse providers. In addition, the program has distributed clinical guidelines in the areas of respiratory care, diagnostic imaging, cardiac care, electrocardiogram and chest x-ray, and clinical laboratory tests.

Use

It is difficult to assess either the totality of protocol-development activity or the extent to which protocols similar to those described above are actually used in ongoing UR programs. There is no common registry for

[1]A number of other medical specialty societies also have similar projects. Sixteen of the 24 organizations who are members of the Council of Medical Specialty Societies responded to our inquiry regarding their development of condition- or procedure-specific criteria. Nine responded that they had developed such criteria for at least some conditions or procedures. Three more stated that they were in the process of doing so. Four responded that they had no such criteria and mentioned no plans to develop them.

such efforts, and many of the protocols developed by commercial firms are not publicly available. Nevertheless, in the course of our study, we had the opportunity to ask many different people in a variety of organizations about the availability of diagnostic-specific protocols.

Many commercial developers of UR products, as well as firms conducting reviews under contract, have developed their own protocols with the assistance of physician consultants—in addition to drawing on protocols developed by others, including the medical specialty societies. For example, Interqual has compiled a manual, SIM III: Surgical Indications Monitoring, that contains over 140 pages of protocols for specific surgical procedures. Another company, Value Health Sciences, is currently developing and testing a computerized system, based on protocols similar to those developed by Rand, that could be used by a broad spectrum of organizations to target individual procedures for appropriateness review. Health Risk Management of Minneapolis, Minnesota, a UR firm, has modified the diagnostic-independent ISD-A criteria into diagnosis-specific criteria for the diagnoses included in the International Classification of Diseases-9th revision, Clinical Modification. Formulated by the firm's staff, the adaptations are also based on a review of the literature and certain medical specialty guidelines.

Health Economics Corporation of Dallas, Texas, uses a somewhat different approach in developing its protocols. Rather than depending on a small group of physicians to develop a protocol for when a procedure is appropriate, they develop norms based on the practice patterns of all physicians in their claims data base. In this way, the potential influence of an individual physician's opinion on the resulting protocol is lessened.

Publications dealing with UR in the private sector, as well as a number of the people we interviewed, indicated that more and more organizations will be using diagnosis-specific criteria in the near future. Insurance companies like Metropolitan Life and John Hancock indicated they were considering such criteria, and Aetna Life Insurance was already testing them. UR firms like HealthCare Compare and Corporate Health Strategies were also considering using such criteria.

For the most part, however, the private sector's main attempt to ensure the appropriateness of individual medical services continues to be voluntary or mandatory second-opinion programs primarily aimed at elective surgical procedures. These programs rely mainly on implicit judgments by physicians regarding the appropriateness of the proposed surgery, rather than on explicit procedure-specific criteria.

Discussion

Diagnostic-specific protocols differ from one another in several ways, including their focus, specificity, and development. The focus of a protocol may be on listing all services that might generally be appropriate for a particular condition or specifying which medical services are appropriate for particular patients. The American Medical Association protocols developed in the early 1970's focused on tests and procedures consistent with a particular diagnosis. More recent protocols, such as those developed by Rand, focus on specifying what medical services are appropriate for patients with specific characteristics; these can be used to detect unnecessary care.

With regard to specificity, among the protocols developed more recently, the individual indications or criteria vary greatly. This is also true of the extent to which key terms are defined and decision rules are stated. Little is known about the extent to which these variations influence the effectiveness of review or the number of cases flagged for appropriateness questions. Logically, it would seem that if the criteria are clinically sound, then the greater the specificity of the protocol, the greater the likelihood of consistently identifying potentially inappropriate care. Greater specificity and clear definitions also decrease the need for medical judgment on the part of reviewer.

Table I.1 is a comparative listing of examples of diagnosis-specific criteria for coronary angiography. As an example of a relatively non-specific criterion that requires considerable judgment, one of the "appropriate" indications in the Blue Cross and Blue Shield Association's protocol for coronary angiography is "coronary artery disease that is suspected or known from history and other clinical data when the test information about coronary anatomy is important for patient management." By contrast, the Rand protocol incorporates greater specificity by separating suspected coronary artery disease into several categories, including chest pain of unknown origin and stable and unstable angina. It also not only requires specific test results (for example, positive exercise electrocardiogram) but defines what a positive exercise electrocardiogram result is.

Table I.1: Examples of Procedure-Specific Criteria for Coronary Angiography

Diagnostic category	Blue Cross/Blue Shield Medical Necessity Program	SIM-III (Surgical Indications Monitoring)	Bay Pacific Health Plan[a]	Hawaii PRO	Rand protocol
Coronary artery disease (not otherwise specified)	Coronary artery disease that is suspected or known from history or other clinical data when the test information about coronary anatomy is important for patient management	Patient in whom the diagnosis of coronary artery disease is suspected but not confirmed by noninvasive means	Abnormalities in treadmill or scintigraphy testing, or wall motion studies	Positive stress test findings (if only indication met, physician advisor must review case)	
Chest pain of unknown origin				Chest pains of uncertain origin when there are (1) documented suspicion of ischemic heart disease in which non-invasive evaluation has failed to provide the firm diagnosis essential for management, and (2) disabling physical symptoms or atypical presentations not explained by comprehensive evaluation	For males, or females over 50, with positive exercise electrocardiograms For females 50 or under with very positive exercise electrocardiograms
Angina (stable or unstable)			Stable or unstable angina pectoris not controlled by acceptable and aggressive medical management Variant angina pectoris	Angina pectoris with documented inadequate response to medical management Documentation that Unstable angina patient has unstable angina or recent onset of angina pectoris	Chronic stable angina and unstable angina—generally only appropriate with positive exercise test results, or failure of medical intervention
Cardiac surgery	Prior to cardiac surgery in patients over 35 with valvular heart disease or in patients of any age with congenital heart disease	Patient for whom cardiac surgery, either coronary artery bypass surgery or other cardiac procedure, is contemplated	Mechanical cardiovascular abnormality requiring cardiac or vascular surgery in the near future Congenital heart disease in adults	Preoperative evaluation of patient with valvular heart disease	Presurgical evaluation of patients for valve surgery (but not appropriate for congenital heart disease)
Other surgery		Patient about to undergo a major non-cardiac procedure in			Preoperative evaluation of candidates for renal

(continued)

Diagnostic category	Blue Cross/Blue Shield Medical Necessity Program	SIM-III (Surgical Indications Monitoring)	Bay Pacific Health Plan[a]	Hawaii PRO	Rand protocol
		whom it is important to assess the status of the coronary arteries			revascularization or abdominal aortic aneurism with either positive exercise electrocardiograms or thallium test, angina, or myocardial infarction
Following coronary artery bypass surgery or percutaneous transluminal angioplasty	Following revascularization procedures if symptoms persist or recur, or signs of altered coronary blood supply become evident	Patient who has undergone previous revascularization, to demonstrate the patency of the grafts		Status after coronary artery bypass surgery or percutaneous transluminal angioplasty with recurrent pain or recent electrocardiogram abnormalities	Following coronary artery bypass surgery with very positive exercise electrocardiogram, angina on mild exertion or on moderate exertion after maximal medical treatment
During acute myocardial infarction				Acute myocardial infarction if complicated by recurrent cardiac insufficiency, severe pump failure, myocardial rupture, mitral regurgitation	During acute myocardial infarction, only with persistent chest pain or a new murmur
After acute myocardial infarction			Post myocardial infarction—patients having a proven coronary event at less than 45 year of age	A young (46 years of age or less) patient following recovery from myocardial infarction (if only indication met, physician advisor must review case)	Depends on type of acute myocardial infarction, presence of angina, and test results
Sudden death survivors			History of near lethal arrhythmias	Resuscitation from cardiac arrest or ventricular fibrillation	Not associated with acute myocardial infarction
Ventricular arrhythmias			Refractory ventricular arrhythmias	Recurrent ventricular arrhythmias refractory to medical therapy	Recurrent ventricular tachycardia with positive exercise electrocardiograms
Other		Pulmonary hypertension of uncertain etiology Symptomatic pericardial heart disease Non-obstructive cardiomyopathy	When non-invasive procedures have not excluded ischemic cardiomyopathy		

[a]Cardiac catheterization indications

Protocol development itself varies in at least two respects. First, some protocols are developed based on an extensive review of the literature on the safety and effectiveness of the particular procedure. The literature review is provided to a group of physicians for their use in developing specific indications for appropriateness. The Rand protocols and Blue Cross and Blue Shield medical necessity protocols are typically developed in this way. In other protocol development processes, such as the American College of Physicians' Clinical Efficacy Assessment Project, the review and development of indications are often combined into one step.

Second, the manner in which the physician panels operate is different. In the Rand process, the panel is given explicit instructions that information on safety and effectiveness is to be given greater weight than "accepted medical practice." Only if the expected benefits of the procedure outweigh the potential risks is an indication considered appropriate. In addition, it is not strictly a consensus process in that the physicians on the panel do not have to agree on the appropriateness of each indication. In most cases, other protocol development processes do, in fact, represent a consensus of a group of physicians, although the number and representativeness of the physicians may not always be clear.

Developmental costs for the protocols are also different, at least in part due to the specificity of the criteria and the differences in approach. The cost of developing and using a Rand-style protocol is estimated to be between $150,000 and $500,000. This cost is considered by many observers to be too high to be practical for the broad range of medical conditions. There is also concern that the state of medical knowledge is not equal to the "rigor" of the Rand approach for many procedures. However, there may be methods of formulating such protocols in a less formal and extensive manner that would still result in criteria that are sufficiently specific to avoid the placing of too great a reliance on the reviewer's judgment.

Diagnosis-Independent Protocols

Diagnosis-independent protocols typically focus on the necessary level of care—including the appropriateness of hospitalization and the need for continued hospitalization—rather than the need for specific proposed treatments. The appropriate level of care is determined with reference to the severity of the patient's illness and the intensity of care

needed for treatment. That is, the issue is whether the patient's condition is (or was) sufficiently serious or the planned care is (or was) intensive enough to warrant admission to a hospital. In addition, after hospitalization, the protocols often can be used to indicate at what point patients no longer require(d) continued hospitalization. The protocols may be either truly generic (that is, diagnosis-independent) or may specify criteria for broad body systems such as the circulatory or nervous systems.

Examples of diagnosis-independent protocols are the Intensity of Service, Severity of Illness, Discharge Screens and Appropriateness (ISD-A) review system, and the Appropriateness Evaluation Protocol (AEP). ISD-A, without the appropriateness component, was developed in 1978. Since then, it has been revised several times, and in 1984 the appropriateness component was added. AEP was developed in the same time period. (See table I.2.)

Table I.2: The ISD-A and AEP Diagnosis-Independent Protocols

Type of criterion	ISD-A criterion	AEP criterion
Severity of illness		
Vital signs	Oral temperature above 103° F (39.4° C)	Persistent fever greater than 37.78° C (100° F) orally or greater than 38.33° C (101° F) rectally for more than five days
	Oral temperature above 102° F (38.9° C) with white blood count above 15,000/cu.mm or bacteria by smear	
	Pulse below 40/minute	Pulse below 50/minute
	Pulse above 200/minute	Pulse above 140/minute
	Respiratory rate above 32/minute	
		Electrocardiogram evidence of acute ischemia; must be suspicion of new myocardial infarction
Blood pressure	Systolic below 80 mm Hg	Systolic below 90 mm Hg
	Systolic above 250 mm Hg	Systolic above 200 mm Hg
	Diastolic above 120 mm Hg	Diastolic above 120 mm Hg or below 60 mm Hg
Laboratory tests	Hemoglobin below 7 grams	
	Hemoglobin above 20 grams	
	Hematocrit below 21%	
	Hematocrit above 60%	
	White blood count above 20,000/cu.mm	
	Serum sodium below 123 mEq/L	Serum sodium below 123 mEq/L
	Serum sodium above 156 mEq/L	Serum sodium above 156 mEq/L
	Serum potassium below 2.5 mEq/L	Serum potassium below 2.5 mEq/L
	Serum potassium above 6.0 mEq/L	Serum potassium above 6.0 mEq/L
	Blood pH below 7.30	Arterial pH below 7.30
	Blood pH above 7.50	Arterial pH above 7.45
	Presence of toxic level of drugs or other chemical substance	
	PaO_2 below 60 torr	
	$PaCO_2$ above 50 torr	
	Culture positive for bacteria or fungi	
		CO_2 combining power (unless chronically abnormal) less than 20 mEq/L or more than 36 mEq/L
Functional impairment (sudden onset)	Sight, hearing, or speech loss	Acute loss of sight or hearing
	Loss of sensation or movement in any body part	Acute loss of ability to move body part
	Extreme weakness without paralysis	
	Impaired breathing	
	Unconsciousness or disorientation	Sudden onset of unconsciousness or disorientation (coma or unresponsiveness)
	Severe, incapacitating pain	
Physical findings	Gross, continuous hemorrhage from any site	Active bleeding

(continued)

Type of criterion	ISD-A criterion	AEP criterion
	Wound disruption (requiring reclosure)	Wound dehiscence or evisceration
	Vomiting/diarrhea with any one of the following: serum sodium above 150 mEq/L; hematocrit above 54%; hemoglobin above 18 grams; urine specific gravity above 1.026; blood urea nitrogen above 35 mg/dL (recent onset); creatinine above 2 mg/dL (recent onset); ileus	
Intensity of service		
Monitoring (at least every two hours)	Special units, vital signs (temperature, pulse, respiration), blood pressure, neurovital signs, urine output, central venous pressure, blood gases, pulmonary artery pressure, arterial lines	vital sign monitoring every two hours or less (may include telemetry or bedside cardiac monitor)
Medications	Intravenous fluid therapy (excluding "keep vein open" and requiring at least 30 ml/kg of body weight in 24 hours)	Intravenous medications and/or fluid replacement (does not include tube feedings)
	Parenteral medications at least 4 times daily	Intramuscular antibiotics at least every 8 hours
	Parenteral analgesics more than 3 times daily	
	Intravenous antibiotic or antifungal agents	
	Intravenous chemotherapy requiring hydration of at least 2000 ml in 24 hours or parenteral medications for control of nausea and vomiting or continuous infusion	Chemotherapeutic agents that require continuous observation for life-threatening toxic reaction
	Initial insulin therapy or insulin pump regulation	
	Volume expanders	
Treatments	Protective isolation Implantation of radioactive materials in doses greater than 30 millicuries	
	Ventilator assistance	Intermittent or continuous respirator use at least every eight hours
Surgical procedure (scheduled within 24 hours)	surgery or procedure not on ambulatory surgery list and requiring general or regional anesthesia (use local ambulatory surgery guidelines)	surgery or procedure that requires either general or regional anesthesia or the use of equipment, facilities, or procedures available only in a hospital

The ISD-A includes a list of approximately 70 generic criteria divided into three categories. The severity of illness criteria include vital signs, laboratory findings, functional impairment, and physical findings. Intensity of services criteria include monitoring, medications, and treatments. Readiness for discharge screens include normal temperature without use of medication, wound healing, and the like. There are also 20 to 60 intensity of service and discharge criteria for each of 12 body systems (for example, cardiovascular and blood systems) that are used if the patient does not meet any of the generic criteria.

Using the ISD-A system, an admission is considered appropriate if a person meets one severity of illness or one intensity of service criterion on the day of admission and one of each by the day after admission. On subsequent days, a day of care is considered inappropriate if none of the intensity of service criteria (generic or body-specific) are met. Similarly,

the patient is considered ready for discharge if all of the discharge criteria are met. There is some opportunity for the reviewer to exercise judgment in cases where several lab values are close to but do not meet individual cutoffs.

The AEP protocol was developed on the basis of a number of previous efforts, including the original ISD admission criteria. The basic AEP is a list of 16 generic admission criteria and 26 day-of-care criteria. If any of 16 criteria is met, the admission is considered appropriate. If any of the 26 day-of-care criteria is met, that day of care is considered appropriate. Unlike ISD-A, AEP allows the reviewer to "override" the criteria and determine an admission or day of care to be appropriate or inappropriate, regardless of the criteria, if there are extenuating patient circumstances. Modifications of the original AEP have been developed for appropriateness reviews in other specific areas, including pediatric, surgical, and psychiatric services.

Another diagnosis-independent protocol, the Standardized Medreview Instrument, was developed based on AEP and ISD-A specifically for a study of inappropriate hospital utilization in the Medicare program. It is longer and more difficult to use than either ISD-A or AEP. In addition, studies have indicated that it does not reliably identify cases of inappropriate admission. As a result, it has not been updated since 1981 and is generally not used.

Indications for Coronary Angiography by Appropriateness Category: Comparison of 1981 and 1987 Samples

Indication[a]	1981 sample[b]	1987 sample[c]
Appropriate		
During evaluation of valve disease	19%	10%
Unstable angina during hospitalization		
Pain controlled	13	20
Pain persists	6	4
Stable angina		
Class III or IV, maximal therapy	9	5
Class III or IV, submaximal therapy, positive or very positive EST	2	2
Class I or II, maximal therapy, positive or very positive EST	2	0
Class I or II, submaximal therapy, positive or very positive EST	4	2
Stable angina following MI, unstable angina, or CABS		
Class III or IV	10	8
Class I or II	2	1
Nonspecific chest pain		
Positive EST and positive thallium or very positive EST	1	1
Positive EST with no thallium	2	[d]
No angina following MI, unstable angina, or CABS, positive or very positive EST	3	1
Asymptomatic, positive, or very positive EST	1	2
During acute MI		
Complicated, no thrombolytic therapy	[e]	4
Uncomplicated, thrombolytic therapy given	[e]	3
Uncomplicated, non-Q wave, thrombolytic therapy not given	[d]	4
Other appropriate	1	1
Total percent appropriate	**74%**	**70%**
Equivocal		
Stable angina		
Class III or IV, submaximal therapy, no or negative EST	3%	3%
Class I or II, maximal therapy, no or negative EST	1	[d]
Class I or II, submaximal therapy, positive EST, positive thallium, no ejection fraction	[e]	1
Class I or II, submaximal therapy, positive EST, no thallium, no ejection fraction	[e]	1
Class I or II, submaximal therapy, no EST, ejection fraction 20 to 49%	[e]	0.5

(continued)

Indication[a]	1981 sample[b]	1987 sample[c]
No angina (asymptomatic, post-MI, or previously stable angina), positive EST, negative or absent thallium	2	1
Nonspecific chest pain, positive EST, no thallium	[d]	1
Stable angina, post-CABS, class III or IV, submaximal therapy	[d]	1
Other equivocal	1	0.5
Total percent equivocal	**8%**	**10%**
Inappropriate		
No angina, no exercise tests performed		
Post-MI, post-CABS, following unstable angina, or following previously stable angina	4%	3%
Asymptomatic or nonspecific chest pain	4	3
No angina, negative EST		
Post-MI, post-CABS, following unstable angina, or following previously stable angina	1	0
Nonspecific chest pain	2	0
No angina, post-CABS, positive EST, no thallium	[d]	0.5
Stable angina, class I or II		
Submaximal therapy, no or negative EST	2	4
Maximal therapy, no EST, ejection fraction unknown	[d]	1
Stable angina, class I or II, post-CABS, submaximal therapy	[d]	2
Congestive heart failure	2	1
During acute, uncomplicated MI	1	[e]
During acute, uncomplicated Q-wave MI, no thrombolytic therapy given	[e]	4
Suspected left ventricular aneurysm without refractory congestive heart failure	[e]	1
Unstable angina as emergency procedure on day of admission to hospital	0.3	1
Other inappropriate	1	0
Total percent inappropriate	**17%**	**20%**

[a]MI = myocardial infarction
EST = exercise stress test
CABS = coronary artery bypass surgery

[b]The sample consists of 1,677 coronary angiography patients. The column may not add to 100 percent because of rounding.

[c]The sample consists of 213 coronary angiography patients. The column may not add to 100 percent because of rounding.

[d]The indication listed was rated in a different appropriateness category by the panel for the sample in this entry.

[e]The panel did not rate the indication as specified but rated a differently defined indication, which is usually listed separately.

Comments From the Department of Health and Human Services

Note: GAO comments supplementing those in the report text appear at the end of this appendix.

DEPARTMENT OF HEALTH & HUMAN SERVICES Office of Inspector General

Washington, D.C. 20201

SEP 7 1989

Ms. Eleanor Chelimsky
Assistant Comptroller General
United States General
Accounting Office
Washington, D.C. 20548

Dear Ms. Chelimsky:

Enclosed are the Department's comments on your draft report, "Medicare: Improvements in the Identification of Inappropriate Hospital Care are Possible." The comments represent the tentative position of the Department and are subject to reevaluation when the final version of this report is received.

The Department appreciates the opportunity to comment on this draft report before its publication.

Sincerely yours,

Bryan Mitchell
For Richard P. Kusserow
Inspector General

Enclosure

Comments of the Department of Health and Human Services
on the General Accounting Office Draft Report,
"Improvements in the Identification of Inappropriate
Hospital Care Are Possible"

Overview

This report focuses exclusively on assessing the appropriateness of hospital care, which GAO defines in terms of decisions about the need for specific services and the proper setting and intensity of those services. In response to a request from the Subcommittee on Health of the House Committee on Ways and Means, this report has three objectives:

-- to critically examine the available information on the extent of inappropriate care in the Medicare program;

-- to describe what is currently being done in both the Medicare program and the private sector to review the appropriateness of medical care; and

-- to suggest approaches that might be effective in reducing the level of inappropriate care in the Medicare program and issues that would have to be addressed if the Health Care Financing Administration (HCFA) were to adopt those alternatives.

GAO found that Utilization and Quality Control Peer Review Organization (PRO) reviews of hospital care have typically identified a lower rate of inappropriate care in the Medicare program than have reviews done by SuperPRO or by researchers. GAO identified a number of differences between PRO and these alternative reviews that might explain the discrepancy. These differences include the criteria used to screen cases, the cases selected for review, and the PRO's reluctance to aggressively question the appropriateness of care. According to GAO, all of the identified differences tend to decrease the rate of inappropriate hospital care uncovered by the PROs.

The primary difference between private sector and PRO review activities is in the timing of reviews. Private sector programs operate prospectively; that is, they focus on the identification, prior to admission, of patients who do not require hospitalization and their diversion to more appropriate health care settings. Except for a small number of surgical procedures that are reviewed prospectively, most PRO reviews are conducted retrospectively; that is, after the patient has been discharged from the hospital. GAO believes that important, and related, differences include the targeting of specific cases for review, the advisory (that is, nonbinding) nature of a determination that the proposed care is inappropriate, and the structuring of the benefits package to include penalties for not obtaining prospective approval and incentives for complying with the review decision.

Page 2

We believe the comparison of findings of inappropriate care by PROs, the SuperPRO, and private entities is questionable. There are too many differences among the three. While the report acknowledges most of its shortcomings, we wish to outline them upfront in our comments. The following points should be considered collectively before criticizing the effectiveness of PRO review, based on the information presented in the study.

— SuperPRO reviewers do not contact physicians, do not understand the demographics of the area, and have lesser workloads. SuperPRO does not know about the availability of alternative settings (e.g., skilled nursing facilities) in the area when making determinations.

— PROs are responsible for reviewing quality of care and must focus a large portion of their retrospective review activities in those areas where quality problems are likely to arise. This, of course, reduces the return on medical necessity and appropriateness reviews.

Additionally, private sector firms used in the study have anywhere from 3 to 5 years experience in utilization review activities versus the PRO/PSRO's experience of 10 years. Also, an analysis of the quality of care and the validation of diagnosis related groups (DRGs) was not included in this report. Finally, only three PROs were interviewed.

We would note that the global decline in Medicare inpatient hospital admission rates from 1983 to 1986 coincided partly with the implementation of the PRO program. This decline in admission rates occurred in a broad range of medical DRGs (Fisher, Health Care Financing Review, Fall 1988), not in a few mandated reviews of surgical DRGs. Whether this decline was induced by PRO reviews or other financial incentives inherent in the Prospective Payment System is uncertain, but, nevertheless, the decline in admission rates perhaps requires some acknowledgment from GAO.

The decline in Medicare hospitalization rates does not compare unfavorably with decreases in hospitalization rates of the non-Medicare population, as shown in data from the American Hospital Association. Thus, it is not clear from general evidence that private sector review methodology is any more or any less effective than PRO review methodology.

It should also be observed that, in general, any judgement on the propriety of a particular medical treatment is associated with a broad range of uncertainty and that nonmedical factors as well as medical factors may justifiably color a particular judgement. A particular operational judgement is likely to err on the side of safety within the broad range of uncertainty. Research judgements or reviews of operational judgements (neither of which have any impact on a person's health) often

do not adequately recognize the breadth of this uncertainty band. Thus, operational judgements are likely to result in higher hospitalization rates than other types of judgements. The GAO study does not sufficiently explore this uncertainty issue.

GAO Recommendation

That the Secretary of HHS require the Administrator of HCFA to expand prospective utilization review in the Medicare program. This should be preceded by a systematic examination of the cost-effectiveness of various options, including those described in this report. If HCFA lacks legislative authority to implement particular options, such as "advisory" prospective determinations and the use of incentives or penalties for complying with prospective review decisions, it should seek that authority from Congress. It is also important to evaluate the current retrospective review requirements in light of the expanded prospective reviews and to avoid unnecessary duplication and improve the efficiency and effectiveness of both types of review.

Department Comment

We agree with the expansion of prospective review. The third PRO Scope of Work expanded preadmission/preprocedure review from 5 to 10 procedures. As we increase our data base, we will be able to evaluate the cost effectiveness of the review and change its focus as needed (e.g., expand review, change procedures reviewed, emphasize review of procedures that can be performed in a lesser setting, etc.). However, we need to balance prospective review with retrospective review in order to assure quality of care.

See comment 1.

In regard to utilizing "advisory" prospective determinations, we do not believe this approach would be appropriate for the Medicare program. We question why advising a physician that a case may not be paid is more of an incentive for finding alternatives than stating that payment cannot be made. In fact, section 1879 of the Social Security Act (i.e., the waiver of liability provision), which affords beneficiaries with financial protection against retroactive denials, demands that beneficiaries receive a notice that unequivocally informs them, where uncovered services are identified, that Medicare payment will not be made. Also, we consider denial of payment to be both a penalty and an incentive to look for alternatives. The incentives mentioned, adjusting deductibles and coinsurance, would not be an option for Medicare, as the statute is currently structured.

Page 4

We agree that it is important to evaluate retrospective review requirements. We do this on an ongoing basis and have changed the review requirements accordingly.

In addition, the third PRO Scope of Work placed greater emphasis on profiling activity by requiring PROs to perform profiles on a variety of data and then adjusting review efforts accordingly. PROs are required to focus on providers with unusual patterns or specific diagnoses or procedures that have been identified as being problematic. For example, in the quality intervention plan, the PRO must profile every physician and hospital and intervene accordingly. Also, the PRO is required to set objectives based upon aberrant practice patterns determined by profiling activity.

Other Matters

Consider implementing a second opinion program.

We have developed a draft Notice of Proposed Rulemaking (NPRM) to implement the second opinion program required by statute. The NPRM is in the final clearance process.

Consider reviewing all elective admissions prospectively.

See comment 2.

Medicare operates within monetary limitations. Prior review of all elective admissions would not be cost effective. In addition, the majority of medical admissions for the elderly are not elective admissions.

Congress should mandate national screening protocols for use by reviewers.

See comment 3.

HCFA has started discussions with a PRO to investigate the possibility of utilizing national criteria, such as the Rand protocols.

Target review on procedures that can be done safely and more cheaply in an outpatient setting.

See comment 4.

The PRO Scope of Work requires PROs to notify the hospitals and physicians as to which surgical procedures can be performed, consistent with accepted medical practice, safely on an outpatient basis. Based upon incentives and exhibited practice patterns in both the public and private sectors, there is already movement of services to the outpatient area.

Page 5

Use the computer to screen cases when conducting prospective review.

See comment 5.

We are in the process of pilot testing the use of a Uniform Clinical Data Set (UCDS). Under the current review system, a typical PRO has a high level of false positives, i.e., cases which fail nurse reviewer screens, are referred to physician advisors for review, but eventually pass physician scrutiny and are approved. The UCDS should reduce false positive referrals because the electronic screening algorithms used in the UCDS can be far more complex and comprehensive than the written criteria used by the PROs. This system will allow a more reliable comparison of review results and statistics among PROs.

Technical Comments

Now on pages 4 and 14.

Pages ES-5 and 1-13

See comment 6.

In regard to the GAO study of coronary angiography, there is not clear consensus in the medical community about the propriety of performing these on an outpatient basis. The controversy becomes even larger when considering aged patients and when considering differences between right- and left-sided catheterizations.

Now on page 4.

Page ES-7

See comment 7.

The report asserts that PROs do not have available penalties against providers other than outright payment denial. This is incorrect; there are penalties other than denial of payment, ranging from intensified review to sanctions and exclusion from the Medicare program.

Now on page 5.

Page ES-8

The report states that: "instead of reviewing individual cases for appropriateness of admissions, the reviews might focus on identifying providers with unusual practice patterns for more intensive review and on verifying information provided during prospective review." The review of the appropriateness of an admission is only one of the seven required reviews performed on each case selected for review. Since this review is performed concurrently with the other six reviews; it basically uses no more resources.

Now on pages 9 and 57.

Pages 1-4 and 4-7

See comment 8.

The report states "Given the extensive experience with prospective review in the private sector". We do not believe that 3 to 5 years experience is extensive.

Page 6

Now on page 11.
See comment 9.

Pages 1-6

The clinical examples given in both paragraphs are flawed. For example, the question: "Does a patient with benign hypertrophy of the prostate gland need immediate surgery?" could be answered positively if the patient has a complete bladder neck obstruction and is unable to urinate at all.

Now on page 12.

Page 1-9

See comment 10.

For quality review, PROs are expected to do more than apply quality screens. They are to screen cases, using the generic quality screens at a minimum, to determine if the care meets acceptable standards of medical care.

In addition, the mandated discharge review focuses on appropriateness in that the PRO determines how long the patient should have remained hospitalized. We admit that this review is mainly to assure that the patient is not prematurely discharged (since the prospective payment system precludes denials for appropriateness in non-outlier cases) but certainly does, philosophically, meet part of the GAO definitions of review for appropriateness.

Now on page 14.
See comment 11.

Page 1-12

"both" should be "each"

Now on page 19.
See comment 12.

Page 2-6

The report states that PROs are not required to conduct any specific review based on profiles. This is not correct. For example, the quality intervention plan in the PRO Scope of Work requires quarterly profiling and outlines what actions the PRO must take as a result of that profiling activity.

PROs are monitored quite closely on their profiling as this is how they identify aberrant patterns by physician, hospital, DRG, etc. Intensification and other interventions are based on results of profiles, as is the development of objectives. Evaluation of PRO performance does include identification of aberrancies by use of profiling.

Now on page 19.
See comment 13.

Page 2-7

PROs also can choose procedures not on the HCFA list, based on supporting documentation that the procedures should be reviewed.

Page 7

Now on page 19.
See comment 14.

Page 2-7

The increase from 5 to 10 procedures subject to preadmission review was instituted by HCFA to promote more effective and efficient review. The congressional mandate for review of 10 procedures is tied to the second opinion program, which has not yet been implemented.

Now on page 19.
See comment 15.

Page 2-8

The statement; "that in the Medicare program, if a hospital stay is determined to be inappropriate, the hospital cannot attempt to collect money from the beneficiary" is not always true; the beneficiary can be liable if an appropriate notice of noncoverage was issued by the hospital and the individual nevertheless requested the services.

Now on page 21.
See comment 16.

Page 2-10

The report states that length of stay is typically not reviewed. Under the PROs' second Scope of Work, length of stay review was done on all speciality cases.

Now on page 23.
See comment 17.

Page 2-13

The list of chart materials reviewed by the PRO is not accurate. The PROs are to review the total medical record (e.g., consultation reports, temperature graphics, notes from specialty units, social service notes, physical therapy notes, radiology reports, etc.).

Now on page 23.
See comment 18.

Page 2-14

If the PRO has illegible copies or not all of the information, the PRO must ask the hospital to resubmit the records.

Now on page 26.
See comment 19.

Page 2-19

We do not approve PRO criteria. We review and recommend changes to the criteria.

Now on page 27.
See comment 20.

Page 2-20

We cannot support the emphasis to focus on utilization problems, as there are many (including members of Congress) that feel the PRO's focus should be on quality. We have attempted in the Scope of Work to achieve a balance between utilization and quality concerns and have structured the reviews accordingly.

Page 8

Now on page 28.

Page 2-21

The Medicare Code Editor is one way PROs target for review specific principal diagnoses that probably represent inappropriate admissions.

Now on page 29.
See comment 21.

Page 2-23 - Note b

"XX" should be 15.

Now on pages 29 and 52.
See comment 22.

Page 2-24 and 3-30

GAO states that unless PROs review for appropriateness of surgical procedures, there will not be a change in the denial rates. It states that only recently have PROs begun review of appropriateness of individual procedures. The review for the appropriateness and the necessity of the admission and the procedure has been a requirement during the first, second and third Scopes of Work.

Now on page 30.
See comment 23.

Page 2-25

One of the PRO's main goals is to ensure the quality of care rendered to Medicare patients. This cannot be done on a preadmission/preprocedure basis. Many categories of review are chosen due to concerns about quality of care or "gaming" of the prospective payment system. In fact, the reviews that are highlighted in the report as having less appropriateness denials than the random sample are those reviews (e.g., readmissions, transfers, specific DRGs, and cost outliers).

Now on page 31.
See comment 24.

Page 2-27

GAO suggests that PROs focus review on bypass surgery, major chest procedures and major bowel procedures. A brief analysis of the second Scope of Work data shows that PROs did review these procedures. Additional analysis of the third Scope of Work data is needed to determine the denial rates in these categories.

Now on page 33.
See comment 25.

Page 2-31

If the PRO denies a procedure as not medically necessary, then payment for the physician's services is automatically denied -- the carrier does not make a determination about those types of cases.

Now on page 35.
See comment 26.

Page 2-34

GAO believes that a simple random sample of cases would yield a higher rate of inappropriate admissions. This is a requirement; PROs review a 3 percent random sample.

Page 9

Now on page 36.
See comment 27.

Page 3-2

SuperPRO not only has no contact with the attending physician, but also does not know the local demographics; thus, for example, it does not know about the availability of alternate sites of care. Also, reviewer workloads are smaller than those of PRO staff.

Now on page 43.
See comment 28.

Table 3.4

This table depicts retrospective and preadmission denial rates. The retrospective rates are correct, however, the preadmission rates do not match the data we have.

Now on page 50 as table 3.7.
See comment 29.

Table 3.8

This table depicts data from PSROs as a basis for inappropriate admissions. The other data used in the report was based on PRO second Scope of Work. The assumptions by GAO are based on second Scope of Work results. The types of review and methodologies vary from PSRO's to PRO's and from second and third Scopes of Work.

Now on page 41.
See comment 30.

Page 3-10

The report assumes medical practice patterns have not changed in 6 years, when, in fact, there have been tremendous advancements in technology since 1983.

Now on page 44.
See comment 31.

Page 3-18

GAO states that "Thirteen States report no preadmission denials while Maryland and Oregon denied 9.8 and 11.1 percent of admissions respectively." These percentages do not match table 3.4. In addition, Maryland was a waiver State and Oregon was under a Corrective Action Plan for data.

Now on page 52.
See comment 32.

Page 3-29

Another explanation for the differences in denial rates is attributable to knowledge of local demographics.

Now on page 55.
See comment 33.

Page 4-3

Prospective review of Medicare patients is far more problematic than the review of non-Medicare patients because of their age and the presence of comorbid conditions.

Page 10

Now on page 57.
See comment 34.

Page 4-7

Contracting with a limited number of firms specializing in utilization review would be contradictory to congressional preference for local review by peers, as many of these companies do not utilize local physicians for review.

Now on page 59.
See comment 35.

Page 4-10

GAO states that HCFA can provide a number of things needed to reconstruct the program according to its recommendations but shows no analysis of what this would entail in terms of resources or time. As the recommendations are major programmatic changes, some analysis of this sort should take place.

Now on page 59.
See comment 36.

Page 4-10

PROs should now have the resources and expertise to perform profiling and the related focusing required by the profile activity.

The following are GAO's comments on the Department of Health and Human Services' letter dated September 7, 1989.

GAO Comments

1. As we explain in chapter 4, we do not agree that advisory determinations are inappropriate in Medicare. Our position is that PRO physicians need incentives to more actively question the care provided by attending physicians. Concerning the question of how to implement a program using advisory opinions, the private sector would certainly be one place to look. Generally speaking, in response to a request for precertification, the reviewing physician would inform the attending physician, as well as the beneficiary, that the proposed procedure or hospital admission could not be certified as appropriate. The notification would go on to indicate that if the proposed procedure or admission occurred, it would be reviewed again retrospectively and a final decision on appropriateness made that would govern whether reimbursement would be made.

There are a variety of ways to carry out the details of advisory opinions that deserve further analysis. Under the current system, the provider would be held liable for the inappropriate care, and the beneficiary would not have to pay. However, with regard to an uncertified hospital admission, the hospital could issue a notice of noncoverage, and the beneficiary might then become liable for any hospital costs associated with the questionable admission. Tying deductibles and copayments to compliance with prospective judgments is another approach that would provide beneficiaries and providers with additional incentives for choosing appropriate medical care.

2. The Department of Health and Human Services states that reviewing all elective admissions would not be cost-effective. However, we continue to believe that this is an open question and recommend that studies be conducted to determine whether the limited information indicating cost-effectiveness in the private sector can be replicated for Medicare. To date, HCFA has provided no evidence of the cost-effectiveness either of its current prospective review requirement or of any potential alternative.

3. HCFA's comment implies that we believe that the Congress should mandate national screening criteria. In fact, our report suggests that HCFA could probably implement national screening criteria without legislative action. Officials of the Department of Health and Human Services report that they are currently using national screening criteria as part of the

Uniform Clinical Data Set. (See also comment 6.) This seems to indicate their agreement with our position.

4. Even though the Scope of Work requires PROs to notify hospitals and physicians regarding procedures performed inappropriately in hospitals, PROs are not explicitly required to question or deny any of those procedures. PROs are also not required to select for review cases admitted to the hospital for those procedures. The Department of Health and Human Services could do more to encourage the movement of services to outpatient settings by instituting more explicit review requirements.

5. We agree with the Department of Health and Human Services that the Uniform Clinical Data Set project is a way of producing more reliable results across PROs, based on electronic screening algorithms, and state this fact explicitly in chapter 4. In addition, in connection with the earlier point regarding national screening criteria, we would like to point out that the type of electronic screening algorithms being developed in the project are, at least implicitly, national standards.

6. The Department of Health and Human Services' comment seems to be related to the first two sentences of the paragraph, which present information on PRO prospective reviews of individual surgical procedures. However, our study of coronary angiography considered only the issue of the need for the procedure, not whether it could have been done on an outpatient basis. We are aware of the controversy over the appropriateness of outpatient cardiac catheterization, but it is not relevant to our point.

7. We believe that the Department of Health and Human Services has misread this portion of the report. We are discussing available options in prospective reviews of individual cases. the Department appears to be referring to "penalties" that are available for care reviewed on a retrospective basis and, typically, across a number of cases from the same provider. In individual cases being prospectively reviewed for admission appropriateness, a payment denial is the only penalty. However, if a provider exceeds a certain number of retrospective denials over a set period of time, PRO can then "intensify" review. Of course, given that the denial rate in those cases is only slightly higher than in the random sample, it is unclear whether the provider really stands to lose much from intensified review. Sanctions and exclusion from the Medicare program occur very infrequently and almost always for quality, rather than appropriateness, problems.

8. While 3 to 5 years may not constitute extensive experience with prospective review, certainly the number of private sector companies doing prospective reviews and the number of hospitalizations reviewed by them are extensive.

9. The Department of Health and Human Services's comment regarding our "clinical examples" suggests that they interpreted the questions as intended to be unequivocal examples of inappropriate care. That was not our intention. Instead, we were simply illustrating the specific types of questions that would be relevant to each of the two appropriateness issues we defined. Consequently, we disagree that the "clinical examples" are flawed.

10. Appropriate changes to the text have been made.

11. Appropriate changes to the text have been made.

12. Appropriate changes to the text have been made.

13. Appropriate changes to the text have been made.

14. Appropriate changes to the text have been made.

15. Appropriate changes to the text have been made.

16. Specialty cases represent only 2 percent of the cases reviewed by PROs under the second Scope of Work. We have, however, changed the draft to acknowledge this fact.

17. The PRO manual suggests that PROs do not have to review the entire medical record; they may, in fact, request partial records, including those portions we mentioned explicitly in the report. We have, however, revised the text to more closely follow the PRO manual.

18. We have revised the draft to make it more technically correct. However, PROs can request the incomplete material only if they know that it was not copied originally.

19. Appropriate changes to the text have been made.

20. Assuming that this particular Department of Health and Human Services comment refers to the types of cases targeted for review, we do not feel that the distribution of cases selected for review (as illustrated

in table 2.1) represents a balance between appropriateness and quality concerns. With the exception of utilization objectives, which are optional under the third PRO Scope of Work, and reviews of specific DRGs, the bulk of the remaining reviews do not contain many inappropriate admissions.

21. Appropriate changes to the text have been made.

22. Our report addresses two broad appropriateness issues: (1) whether an individual patient needs or needed a particular service and (2) the appropriateness of providing needed services in a hospital setting. As HCFA states, PROs have been expected to conduct reviews of both issues under the first and second Scopes of Work. However, the addition of an explicit requirement in the third Scope of Work to examine the necessity of particular invasive services, as well as our interviews with three PRO and HCFA officials, suggests that, in practice, little review of the need for surgical procedures was going on. (Most reviews of appropriateness focused on whether the hospital was the appropriate level of care.)

23. The Department of Health and Human Services' comment is quite correct. In fact, it is simply a restatement of our argument that PROs do not focus on appropriateness issues and, as a result, do not do as much as they could to identify and deal with inappropriate care in the Medicare program. (We are not, however, suggesting that quality review is unnecessary.)

24. Our primary point in this section of the report is that the Department of Health and Human Services (through PROs) could better target reviews toward particular categories of cases with high levels of inappropriate admissions, based on prior research. HCFA has singled out only three categories of procedures from a more extensive list that we present of the types of cases frequently found to be inappropriate admissions. With regard to what PROs did under the second Scope of Work, we present information in table 3.1 on procedures reviewed prospectively by PROs. Coronary artery bypass surgery was reviewed by only six PROs, and the other two categories by fewer than five PROs. (The truth of the latter assertion follows from the fact that the table excludes such cases.)

25. In the case of the unnecessary surgical procedures cited by HCFA, as well as of unnecessary physician services associated with inappropriate cost outliers, PRO must notify the carrier, and the denial is then automatic. However, in all other cases, the fiscal intermediary notifies the

carrier, and the carrier must then make an independent assessment. The text of our report has been modified to acknowledge these situations.

26. We have made some minor changes to the text to clarify our point. We are not suggesting that PROs should review a random sample of cases in order to obtain a yield of a higher proportion of inappropriate admissions. Rather, our point is that the rate of admission denials in the cases from the 3-percent random sample is higher than in most of the other categories of cases reviewed. In view of this disparity in admission-denial rates, we believe that HCFA and PROs could do a better job of targeting categories of cases with high rates of inappropriate admission than they currently do. This conclusion is further supported by the higher rate of admission denials in those cases selected for intensified review based on a pattern of inappropriate care identified by PROs.

27. One can understand the importance of being aware of alternative sites of care in making decisions about the appropriateness of hospital admissions. However, if PROs continue to accept cases of questionable appropriateness based on these reasons, local communities will feel no pressure to develop the necessary resources. Further, there may already exist appropriate nonhospital care settings that could be used if providers had incentives to identify them. With regard to the Department of Health and Human Services' suggestion that smaller workloads increase the likelihood that truly inappropriate care will be identified, the Department may want to reconsider the resources it provides to PROs to conduct their reviews.

28. Our calculations of preadmission denial rates in table 3.4 are based on information taken from the cited HCFA report on pacemaker reviews (page B-5), DRG and procedure reviews (page B-6), and other preadmission reviews (page B-7). We did this to focus attention on preadmission/preprocedure reviews. The calculations have been rechecked. The relevant footnote to the table has also been expanded to make this point clearer.

29. The data in table 3.7 do not represent Professional Standards Review Organization (PSRO) data. In half of the studies, the data are from special studies of inappropriate admissions conducted with the assistance of PSROs. In the other studies, PSROs were not involved at all. In fact, the final study in table 3.8 was conducted after PROs were in operation. The remaining points are correct and have been acknowledged a number of times throughout the report. Nevertheless, we would like to reiterate that the pattern of results across review approaches and

over time is so strong that we believe the only reasonable conclusion is that PROs could do their appropriateness reviews more effectively.

30. The report makes no such assumption. In fact, in recognition of the fact that medical practice has changed, we contracted with Value Health Sciences to conduct a limited replication of the earlier Rand work using updated criteria and more recent Medicare records. Despite the changes in medical practice over time, the results from the two studies are very similar.

31. Appropriate changes have been made to the text. In particular, we obtained corrected data for Oregon from HCFA.

32. See comment 27 concerning "local demographics."

33. We recognize the fact that Medicare patients are older and have more comorbidities but question the assertion that this makes prospective review "far more problematic."

34. We agree that the Congress, in the past, has expressed a preference for local review by peers. In fact, in the report we suggest that the Department of Health and Human Services may want to request specific legislative authority for moving away from "local" control of Medicare reviews. At the same time, PROs are currently state-wide rather than strictly "local," and, in some areas, out-of-state organizations conduct the PRO functions. In addition, the Department has already indicated in its comments regarding national criteria and the Uniform Clinical Data Set project that it is moving away from a strict local-determination policy. We view the option of contracting with a limited number of utilization review firms as an extension of an ongoing loosening of the "local control" tradition, which may result in more effective and consistent reviews.

35. We chose not to attempt to estimate the specific time and resources because we are recommending that HCFA examine a variety of options. If we had proposed a specific solution, a more detailed analysis of costs would have been appropriate.

36. The Department's comment refers to current resources for profiling activities. Our point concerns the need to consider the resource implications of an expanded and redirected set of profiling activities.

Major Contributors to This Report

Program Evaluation and Methodology Division

Linda K. Demlo, Assistant Director
Roger B. Straw, Project Manager
Michael Hutner, Evaluator

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are $2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100