MECHANISMS FOR ASSURING QUALITY OF U.S. MEDICAL CARE SERVICES: PAST, PRESENT, AND FUTURE

PREPARED UNDER A GRANT FROM THE DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

ROBERT H. BROOK,
ALLYSON DAVIES-AVERY

R-1939-HEW
AUGUST 1977
The research reported herein was performed pursuant to the Health Insurance Study grant from the U.S. Department of Health, Education, and Welfare, Washington, D.C. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any agency of the United States Government.
PREFACE

A significant feature of the experimental portion of the Rand Health Insurance Study will be an assessment of the quality of health care delivered to study participants. This report reviews the variety of approaches to quality assessment and assurance taken by the private and public sectors. It describes ongoing quality assurance programs and analyzes their potential for improving the quality of medical care and thereby the health of the American people. Equally important from a research perspective, the report identifies and discusses problems that require additional research, with the aim of refining current methods of quality assessment and developing new ones. The report was prepared with support from the Rand Health Insurance Study grant (016B-7501-P2021) from the Department of Health, Education, and Welfare.

An earlier and substantially different version of this report entitled "Quality Assurance Mechanisms in the United States: From There to Where?" appeared as a chapter in A Question of Quality--Roads to Assurance in Medical Care, edited by Gordon McLachlan and published for the Nuffield Provincial Hospitals Trust by the Oxford University Press in 1976.
SUMMARY

In the past few years, the United States has been propelled into the forefront of the movement to develop programs for assuring the quality of medical services. The reason lies chiefly in three recent events. Escalating costs in the federally funded Medicare and Medicaid programs led to passage of P.L. 92-603, which contains provisions for establishing Professional Standards Review Organizations. Disenchantment with the efficiency of the prevailing fee-for-service delivery system led to passage of the Health Maintenance Organization Act of 1973, the intent being to encourage establishment of prepaid group practice systems, which are required by law to establish their own quality assurance systems. And escalating costs of malpractice insurance coverage led to the "malpractice premium crisis" of the mid-1970s; as a result, many state legislatures have mandated quality assurance activities while considering alternatives to physician liability through the tort system.

The quality assessment field has a far longer history than these recent events would suggest. Attempts to measure the quality of medical care have gone on for over a century, although they did not gain prominence in the United States until two decades ago. During the past twenty years, the research focus has shifted from the structural attributes of the medical care system to the technical processes of care and, more recently, to assessment of health outcomes for patients. Most of these studies have documented deficiencies in the quality of care, some of them serious, but the results have seldom been used to improve the quality of services. Furthermore, there is little agreement over which basis for quality assessment—structure, process, or outcome—produces the most valid and reliable judgment.

Within the past decade, the private and public sectors have mounted several quality assurance programs, which have taken a variety of approaches to the task. The PSRO programs are organized and operated by physicians at the local or area-wide level; their
current focus is on the quality of hospital services and cost-control, using explicit process criteria. The Performance Evaluation Procedure was developed by the Joint Commission on Accreditation of Hospitals, a voluntary body; its review procedure focuses on both the process and end-results of hospital care for individual patients. Unlike these programs, those mandated by the Health Maintenance Organization Act for all federally funded HMOs are required to assess both process and outcomes and ambulatory care as well as hospital services.

The malpractice premium crisis has impelled other steps. Many state legislatures view the threat of malpractice suits as a type of quality assurance. Consequently, they have been reluctant to change the "current malpractice system" without first attempting to guarantee improvements in the quality of care by altering other aspects of the medical care system. They have done so by requiring physicians to undergo relicensure, recertification by specialty boards, and mandatory continuing education. There is little objective evidence, however, to support the assumption that such actions will improve the quality of care.

Analysis of the virtues and defects of the current approaches to quality assurance points clearly to the need for further research:

- Research is needed to test the usefulness of quality assessment in the ambulatory care area; eventually, quality assurance programs must integrate review of both inpatient and outpatient care.
- The current practices of local selection of health problems for peer review and local establishment of criteria and standards must be tested against two assumptions for which empirical evidence again is lacking: that physicians have the necessary epidemiologic information to develop such standards, and that physicians are more apt to change their practice behavior in response to locally determined standards than they are to national standards.
- The quality of care throughout an institution or region is often judged by the care given to patients with only one or two indicator conditions; the validity of such generalization must be tested rigorously.
-vii-

- Extensive study is called for to resolve the issue of which type of data (structure, process, or outcome) and which data source (medical record, claims form, or direct observation) produces the most reliable and valid basis for quality assessment.
- Most quality assessment studies have dwelt on the technical aspects of care, ignoring the art-of-care component; quality assurance systems must incorporate efforts to improve both aspects of care.
- Performance of the quality assurance systems themselves must be reviewed and evaluated to determine their effectiveness in changing physician behavior and the health of the people.

This is not to say that current programs should be delayed until they can be polished to near-perfection. Present-day assessment methods are able to detect deficiencies in the quality of care. Inaction is the greatest threat to the continuation of research and to the ultimate success of quality assurance systems—the failure to use information from quality assessment studies to improve medical care.
ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions made by Robert Gladstone of The Rand Corporation and Charles E. Lewis of the School of Medicine at the University of California, Los Angeles, in reviewing this report. Their comments and critiques were helpful in clarifying certain points of the argument for readers not closely familiar with the quality assurance mechanisms and policy issues discussed. The authors also thank Joseph P. Newhouse of Rand for his general guidance and warm support of this endeavor. Responsibility for any errors in interpretation rests solely with the authors.
CONTENTS

PREFACE .............................................................................................................. iii
SUMMARY ........................................................................................................... v
ACKNOWLEDGMENTS ......................................................................................... ix

Section
I. INTRODUCTION ............................................................................................. 1
II. A BRIEF HISTORY OF QUALITY ASSESSMENT ............................................ 3
III. CURRENT QUALITY ASSURANCE SYSTEMS .................................................. 10
    Professional Standards Review Organizations ............................................ 10
    Performance Evaluation Procedure ............................................................. 17
    HMO Quality Assurance ................................................................................ 20
    Crisis in Malpractice Premiums .................................................................... 22
    Concluding Observations .............................................................................. 23
IV. IMPLICATIONS FOR CONTINUING RESEARCH IN QUALITY ASSURANCE ................................. 25
    Hospital Versus Ambulatory Care Review ................................................... 25
    Local Control ................................................................................................. 27
    Scope of Review Activities ........................................................................... 29
    Type and Source of Data ............................................................................... 29
    Technique Versus Art of Medical Care ......................................................... 36
    Evaluation of Quality Assurance Systems .................................................. 36
    Duplication of Systems ................................................................................. 37
V. CONCLUSIONS ................................................................................................. 38

REFERENCES ....................................................................................................... 39
I. INTRODUCTION

Some critics have described the American medical care system as an amorphous "nonsystem" employing antiquated organizational techniques unsuited to an industry that consumes more than $100 billion a year (White, 1968). Critics find a variety of faults with the system: no evident overall objective; lack of planning; lack of a national financing mechanism; lack of coordinated delivery systems with responsibility for defined populations; and inequity of access to medical care resources. It is debatable whether all of these criticisms are justified. Other countries have responded to similar criticisms by organizing fundamentally new or even radical (by United States standards) medical care delivery systems; it is not certain, however, that the changes have produced improvements in health (Townsend, 1974).

The reason for the uncertainty is that none of those countries simultaneously mounted programs to measure and then, if necessary, improve the quality of medical care. Such programs are obviously needed, given the marked shift in the primary function of the personal medical care system over the past century. A hundred years ago, that function was to provide compassionate care for patients. Today, advances in biomedical sciences have added another function: efficient delivery, to the entire population, of medical services that result in cure or control of disease and in maintenance or improvement of health. Some type of quality assurance system appears necessary to measure the performance of this second, and apparently predominant, function.

The United States finds itself pushed to the forefront in developing such a system, largely owing to three forces:

- Escalating costs in the federally funded Medicare and Medicaid programs led to passage of P.L. 92–603 (1972), which contains provisions for establishing Professional Standards Review Organizations.
- Disenchantment with the efficiency of the prevailing fee-for-service delivery system led to passage of the Health Maintenance
Organization Act of 1973 encouraging establishment of prepaid group practice systems, which are required by law to establish their own quality assurance systems (P.L. 93-222, 1973).

- Escalating costs of malpractice insurance coverage led to the "malpractice premium crisis" of the mid-1970s; as a result, several state legislatures have mandated quality assurance activities (such as relicensure, required continuing education, and state boards of quality assurance) as they consider alternatives to physician liability through the tort system (Brook, Brutoco, and Williams, 1976).

The purpose of this report is threefold: 1) to place in historical perspective recent developments in quality assurance in the United States; 2) to describe current quality assurance programs and analyze their potential for improving the quality of medical care and thereby the health of the American people; and 3) to indicate where continuing research is necessary on the methods of quality assessment and on improving the level of care rendered.
II. A BRIEF HISTORY OF QUALITY ASSESSMENT

It would be pedantic to begin this review of efforts to assure quality of care with examples taken from the ancient Greeks and Egyptians or from the Bible, although such efforts are as old as the practice of medicine itself. It will be more illuminating to begin with several ideas popularly supposed to have developed during the last two decades in the United States, which actually originated over a century ago in the United Kingdom.

During the Crimean War, Florence Nightingale (1858) conducted a series of studies of the quality of hospital care available to the British Army. She used such data as the number of hospital deaths by diagnostic category to indicate the effects of unsafe conditions prevailing in Army hospitals, and argued that improvements in sanitary conditions could dramatically reduce these case-fatality rates.

A few years later, Nightingale (1863) proposed what could be considered the rudiments of a uniform hospital discharge abstract system. Its objective was to relate the use of hospital beds to indicators of health in order to encourage the most efficient and effective use of beds. Among the data she recommended for collection were the number of patients admitted to the hospital during the year, the number who died in the hospital and the number who recovered and were discharged, and the number discharged as incurable or unrelieved, or who left the hospital at their own request.

Nightingale recognized that analysis of such data, after controlling for demographic variables, would provide valuable information for both treatment and hospital administration. Specific medical and surgical treatments could be correlated with diagnostic categories; mortality rates could be calculated for each diagnostic category; and the proportion of hospital beds used for restoration of function, as opposed to providing a place to die, could be determined. Control for patient case-mix was probably unnecessary for this level of analysis, because the major purpose was to identify gross deficiencies in care and major inefficiencies in use of resources.
Florence Nightingale, then, appreciated the need to examine both what is done to and for patients and what happens to them afterward—both the process and the outcome of care. (The importance of such analyses has recently been recognized in the United States [Murnaghan and White, 1970]. A demonstration project of the Uniform Hospital Discharge Data Abstract has been completed [Hodgson, Kucken, and Ensign, 1973], and the system is now being adopted by hospitals nationwide.)

Half a century later, Groves (1908) issued a plea for uniform registration of the results of operations, arguing that:

If . . . a surgeon makes a specialty of some disease or operation and tabulates all his own results, or another by chance has some notable successes and records them, or the author of a textbook collects published records of various writers and summarizes them, is it not obvious that such collection of figures will represent the best and not the average results?

To obtain information about "average results," Groves conducted a survey of the 50 hospitals in Great Britain with over 200 beds. Data from the 27 hospitals responding showed a 44 percent operative mortality rate from radical operations for malignant diseases of the stomach, a 24 percent mortality rate from prostatectomy, and a 9 percent mortality rate from appendectomy.

Groves' proposal for uniform registration raised two generic points: the need to develop a standard classification for diseases and operations that would permit comparisons of data from different hospitals; and the need to establish a follow-up system for particular categories of disease, such as malignancies, that would allow assessment of long-term results, such as mortality, degree of disability, or level of symptoms.

Only a few years later, Codman (1914), a surgeon at Massachusetts General Hospital, lamented the lack of outcome assessment in the United States:

One might say that the instruction of the students is irrespective of the results to the patients, but let us
suppose, in surgery, for example, that all the operations which have been watched by these students have been mis-directed efforts at the cure of disease, and the students have learned to do something which is not worthwhile and does not really improve the patient. The product of the hospital in this case, even as regards student instruction, would be nil—even worse than nil. We are, therefore, referred again to the classification of disease and the results to the patients, because a student would naturally wish to receive his instruction at a hospital where the treatment was shown to be of benefit to the patient. We may then say that the product of the hospital in medical education, like the product in the number of cases treated, depends on whether or not the cases are well treated.

To determine whether patients were well treated, Codman attempted to institute a patient follow-up system at Massachusetts General. His objective was to raise his own level of performance by examining, one year later, all patients he had operated on. He thereby hoped to determine whether the operations had been indicated and had benefited the patients.

Thoroughly frustrated in his initial effort, Codman resigned his position and started his own hospital where he instituted his follow-up system. Each surgical patient was recalled a year after operation for assessment of health status in terms of the original objectives of the operation. Codman was then able to determine whether his original diagnosis was correct, whether the operation was a technical success, and whether the patient had benefited or whether the operation had produced some untoward or iatrogenic effect.

Considering, in retrospect, the significant contributions that Nightingale, Groves, and Codman had made by 1914, it is disappointing that little substantive work in quality assurance went on in the United States during the next three decades. Perhaps the Flexner Report (1910), which prompted major improvements in the structure and content of medical education in the United States, had an untoward result as well. Since medical education was presumed to be adequate after Flexnerian reforms, it may have been considered unnecessary to measure the results of care delivered by physicians trained in the new schools.
When interest in assessing the quality of care reawakened in the late 1940s and 1950s, its focus underwent a striking change. No longer was quality assessment based on end-results of care. Instead, the emphasis was on the adequacy of diagnostic investigations and therapeutic interventions—the process of medical care.

Three landmark studies were published during this period: the results of a 1953–1954 study by Peterson et al. (1956) of the quality of care delivered by general practitioners in North Carolina; Morehead's (1958) study of the quality of ambulatory care provided in the Health Insurance Plan of New York City (a prepaid group practice); and Fitzpatrick, Reidel, and Payne's study (1962) of the quality of care rendered in a select group of short-stay general hospitals in Michigan. Peterson et al. observed general practitioners while they were providing care, and scored their practice on the basis of adequacy of history, physical examination, therapy, and type and amount of follow-up care. Morehead's assessment relied on physicians' implicit judgments of the quality of the process of care, which they arrived at both by reviewing medical records and by talking to physicians who provided the care. Fitzpatrick et al. judged adequacy of the process of care by comparing the information contained in medical records against a set of explicit, disease-specific criteria established a priori by a group of physicians. All three studies found major deficiencies in the care provided, suggesting that many practicing physicians took incomplete histories and performed inadequate physical examinations.

Other attempts to assess quality of care during this period focused on structural variables: innate characteristics of physicians, such as age or length of training, and of facilities, such as soundness and safety of the physical plant, or staffing patterns. The best-known proponent of this type of assessment was the Joint Commission on the Accreditation of Hospitals, which sent expert teams to hospitals to evaluate their quality against a checklist of minimum standards.

The shift away from outcome assessment as an indicator of quality of care during the 1940s and 1950s may have been motivated in part by
concerns for practicality and feasibility. Information about what a physician did to and for the patient could be readily obtained from the medical record. Information about the end-results of care could be obtained only from follow-up patient interviews, an expensive procedure that requires extensive cooperation from physicians and patients. At the same time, the medical community was excited by new procedures that used sophisticated instruments to probe the inner recesses of the body. This uncritical fascination with technology did not encourage questioning the value of these procedures. This changing view of medicine, emphasizing use of advanced biomedical technology rather than "laying on of hands," may also have encouraged the new focus on process instead of end-results.

Since 1965, a combination of circumstances had prompted renewed efforts to assess quality of care by measuring patient end-results. Costs of personal medical care services have risen rapidly in all countries in the face of increased demands by all social systems (such as education) on scarce resources. In the United States, the federal government has tripled its contribution to financing medical care in the past decade, largely through the Medicare and Medicaid programs, which have been major factors in the cost increases; it has also intensified regulation of both costs and quality of services purchased through federally funded programs. Finally, there is a growing concern that improvement of financing and delivery of medical care services may be less effective in maintaining or improving health than would improvement of other social conditions, such as housing or employment.

The need to control costs, and increased questioning of the effectiveness of medical care practices, instigated a variety of efforts to evaluate the efficacy and effectiveness of expensive but commonly accepted procedures in terms of their impact on patient outcomes. To date, most of these studies have been undertaken in the United Kingdom. For example, Mather et al. (1971) performed a randomized clinical trial to examine the relative efficacy of the coronary care unit versus home care in treating people who have had acute heart attacks. The study found that, for groups of patients
assigned randomly to home or hospital care, hospitalization in the coronary care unit did not decrease the risk of death. In another clinical trial, Weddell and his colleagues (Chant, Jones, and Weddell, 1972; Piachaud and Weddell, 1972) found that outpatient treatment of patients with varicose veins using an injection/compression procedure produced an equally satisfactory cosmetic result and less morbidity than did the generally accepted and more costly inpatient operation. Finally, an experiment performed in the United States concluded that hypoglycemic agents, the usual treatment for patients with adult-onset asymptomatic diabetes, may increase mortality instead of decreasing it (University Group Diabetes Program, 1970).

If the results of evaluations of medicine as it was practiced twenty years ago are still valid, and if deficiencies persist in such basic areas of practice as physical examination and history-taking, then there is reason for concern about the way in which physicians use today's sophisticated drugs and therapies—for which the potential for harm may rival the potential for benefit under ideal circumstances. Many powerful new therapies—antibiotics, intensive care units, radical surgical procedures, antineoplastic drugs—save lives but also may produce serious iatrogenic disease. Their effectiveness depends on the precision and skill with which they are applied. Given this realization and an environment that permits the value of technical medicine to be questioned, establishment of a quality assurance mechanism for the personal medical care system is a practical necessity that could become a reality.

The foregoing review brings us to a description of the current state of the art of quality assessment. Interest in evaluating quality is now common to medical care systems worldwide, regardless of their particular form of financing or organization; it is not a concern solely of fee-for-service systems. Efforts to evaluate quality have gone on for over a century, during which time research has documented major deficiencies in quality of care. Documentation has not generated effective action, however. Most of these studies have been merely descriptive assessments of the quality of care, and few of their results have been fed back to the respective programs or practices for use in improving quality of care.
No new conceptual frameworks for assessing quality have been developed during the last two decades. Although three time-honored approaches stand out—assessment of quality using structural, process, and outcome variables—there is no evidence to indicate which one produces the most valid judgments of the program or practice being assessed. In each approach judgments of quality have been based on either implicit or explicit criteria. Here again, opinion differs as to which type of judgment is the most valid. Data on quality of care have been obtained from a variety of sources, ranging from claims for payment of services rendered to medical records to direct observation of medical practice. It is not known to what extent the data source used determines the validity and reliability of the quality assessment. (A more thorough discussion of these issues is presented elsewhere; see Brook and Avery, 1976.)
III. CURRENT QUALITY ASSURANCE SYSTEMS

This section considers current policies and programs for assuring quality of personal medical care services in the United States. Each has contributed a new acronym—PSRO, PEP, MAP, and TAP—to the already bewildering array of shorthand in the medical care field. It is easy to make light of this "attack of the alphabet," but the programs they stand for deserve serious consideration. The variety of proposed systems and their approaches and the variety of sponsoring agencies (both public and voluntary private bodies) are indicative of the complex and pluralistic nature of the American medical care system. As these programs become fully operational, efforts to coordinate their activities and avoid duplication will become necessary.

PROFESSIONAL STANDARDS REVIEW ORGANIZATIONS

The best-financed—and most controversial—of the quality assurance systems in the United States is the Professional Standards Review Organization Program (PSRO). It was established in 1972 in an amendment to the Social Security Act. The program is operated at the federal level by the Health Care Financing Administration in the U. S. Department of Health, Education, and Welfare.* The legislation requires the federal government to assure that its financial resources are used efficiently and that services are of high quality. The law (P.L. 92-603, 1972) provides that: 1) payment for services covered under certain sections of the Social Security Act (Medicare, Medicaid, and Title V [Maternal and Child Health]) programs will be made only when and to the extent services are judged medically necessary; 2) payment will be made for services provided by a hospital on an inpatient basis only when such services cannot be provided effectively on an outpatient basis, or more economically in an inpatient health care facility of a different type; and 3) PSROs are responsible for determining that services and items are medically necessary and that the quality of such services meets professionally recognized standards of medical care. To accomplish these tasks, the PSROs have been charged with three functions: utilization review

(focused chiefly on cost containment); medical care evaluation studies (the major quality assurance effort); and profile analysis, or retrospective review of patterns of care.

At the present time, PSRO's scope as a quality assurance system is limited in two important ways. Only the care delivered to persons receiving care through federally financed programs (chiefly Medicare and Medicaid) will be reviewed, and only services rendered in an institutional setting—the hospital and nursing home component of medical care—are currently subject to review. (The law does provide for eventual review of ambulatory care delivered to participants in the above-mentioned programs. Although several demonstration projects involving PSRO review of ambulatory care are in operation, * it is unlikely that full implementation of the ambulatory care review phase of PSRO activities will occur soon.) Medicare covers part of the care provided to low-income persons. Thus, the PSRO program will monitor costs and quality of care delivered to approximately one-fourth of the United States population, which consumes about one-third of its medical care services.

It is interesting to consider why a law was passed that regulates only one part of the medical care industry. The Medicare and Medicaid programs began operation in the mid-1960s. By the early 1970s, costs, particularly for hospital services, had so far exceeded budgetary expectations that attempts at cost control were inevitable. Narrowing the scope of benefits was politically unattractive at the federal level, which has responsibility for virtually all Medicare financing. (Certain cost-sharing provisions necessitate out-of-pocket expenditures by enrollees.) Several states, however, have narrowed the scope of benefits in their Medicaid programs, which are jointly funded by state and federal governments. Changes in cost-sharing provisions for Medicare enrollees were made at the federal level, chiefly by increas-

* Approximately six ambulatory care review demonstration projects were funded in 1976 for a two-year period in PSROs in New Mexico, Multnomah County (Oregon), North Bay (San Francisco), Bronx (New York), South-Central Pennsylvania (Harrisburg), and Central Massachusetts (Williams, 1976).
ing the amount of the annual deductible. Although the original legislation allowed increases in the annual deductible to keep pace with increasing medical care prices, the higher out-of-pocket expenditures these increases implied were also politically unattractive. Utilization review committees were set up in participating hospitals and charged with reviewing the necessity for Medicare and Medicaid hospitalizations, but they were ineffective in stemming the rising use of services and, with it, rising costs to the federal government.

During the same period, peer review activities were developing within the profession. The "medical care foundation" movement, which traces its roots to California in the mid-1950s, gained considerably in numbers during the early 1970s. These foundations are not-for-profit organizations of a few hundred to a few thousand physicians who practice in the same area. While the functions of individual foundations vary considerably, the "foundation" is best known for its review of the ambulatory and hospital care delivered by participating physicians to determine appropriateness and quality of services before authorizing payment by fiscal intermediaries. (For a lengthier discussion of the foundation movement, see Egdahl, 1973, Sasuly and Hopkins, 1967, and Steinwald, 1971.)

This interest in peer review resulted in the participation of several foundations in the Experimental Medical Care Review Organization (EMCRO) program (National Center for Health Services Research and Development and Arthur D. Little, Inc., 1973; Sanazaro et al., 1972; and Goldstein et al., 1975). The dual purpose of this federally funded demonstration project was to determine if medical foundations, organized as they were on an area-wide basis, could improve quality of care and increase the efficiency with which scarce resources were used. (Not all EMCROs were medical care foundations and not all engaged in peer review activities; for example, the EMCRO established at the University of California, Los Angeles, was engaged chiefly in developing sets of process criteria for medical audit based on a branching logic or decision-tree approach; for an example of their work, see Greenfield, et al., 1975.)

Some leaders of the foundation EMCROs were quick to argue that the review activities stabilized costs and improved quality of care.
Objective supporting evidence is limited, however. A few articles (reviewed in Brook and Williams, 1976) suggest that one or two of these programs contained costs by decreasing hospital use. Because of weak evaluative techniques and the relatively small decreases found, however, their conclusions are extremely controversial and may be erroneous. A recent evaluation of the New Mexico EMCR0’s first two years of operation found no effect on use of hospital services for EMCR0’s preadmission certification and recertification programs, which determined medical necessity for hospitalization for Medicare and Medicaid patients (Brook and Williams, 1976). *(Whether this EMCR0’s effectiveness has increased in recent years is now being evaluated.)*

Despite the lack of solid objective evidence of influence on costs and quality of care, it seemed reasonable from both a political and practical point of view to model PSROs on the medical care foundations and EMCR0s. These organizations were area-wide and open to all physicians, many physicians supported them, some information documented their effectiveness, and rhetoric suggested that cost-containment and quality improvement were possible. The needs of the federal government and at least some of the medical profession could conceivably be met by development of PSROs.

**Implementation**

Implementation of the PSRO program began in earnest after January 1974, by which time 203 PSRO areas had been designed. The promise of $50,000 in planning funds, and the implied threat that the Secretary of Health, Education, and Welfare (DHEW) could designate any group, including a nonphysician group, as the PSRO if local physicians failed to organize a PSRO before July 1976, prompted physicians to begin

---

*This organization, however, did have a marked effect on one aspect of quality: the inappropriate use of injectables. Education of physicians and explicit denial of payment for claims submitted for inappropriate injections decreased the number of injections from 41 per 100 ambulatory visits billed to 16 per 100 visits billed (a 60 percent reduction) over the two-year period. Because half of these injections were antibiotics, of which 75 percent were judged by peer review physicians to be medically unnecessary, iatrogenic complications of medical therapy were obviously prevented.*
forming local groups. A recent amendment to PSRO legislation gives
those areas in which there is considerable resistance to formal peer
review (such as Louisiana, Nebraska, and Texas) until December 1978
to organize a physician-sponsored PSRO, after which time the Secretary
of DHEW can designate any group (medical school, health department, or
fiscal intermediary, for example) as the local PSRO (Williams, 1976).

Each area-wide PSRO is a not-for-profit corporation in which all
licensed physicians, both medical doctors and osteopaths, are eligible
for membership. Membership is voluntary, and the PSRO cannot charge
membership dues. If more than 10 percent of the physicians in a PSRO
area object to the particular group organized, the law requires a vote
of all physicians practicing in the area to accept or reject the pro-
posed PSRO. Once such an organization is founded, it competes through
a grant process for federal planning funds. After successful comple-
tion of the planning stage (usually one to two years), the PSRO may
obtain conditional status and in another year or two becomes fully
operational. It is then fully responsible for determining that quality
assurance and utilization review activities meet minimum standards in
all hospitals serving Medicare, Medicaid, and Title V patients in its
area.

As of late summer 1976, some 20 PSROs were in the planning stage
and 100 had conditional status; none were fully operational. Another
60 groups in areas without PSROs had filed "expressions of interest"
and were expected to receive planning funds in the next fiscal year.
In remaining areas, either no local or statewide groups had expressed
interest in organizing a PSRO, or the area (entire states, in several
cases) had expressly indicated no interest in participating in the
PSRO program (Williams, 1976). The 100 conditional PSROs (for which
such data are available) are responsible for quality assurance in from
3 to 149 hospitals each; the total population covered by individual
PSRO areas is estimated to range from approximately 260,000 to
3,400,000.

* Personal communication, Bureau of Quality Assurance, Health
Services Administration, Department of Health, Education, and Welfare,
November 1976.
PSRO Quality of Care Assurance

To carry out its legal mandate, each PSRO will be responsible for seeing that studies directed at utilization review and quality assurance are performed in each of the area hospitals. The PSRO may delegate authority to perform these studies to each hospital's utilization review committee, but the PSRO remains responsible for fulfilling legal requirements. The utilization review activities of the PSRO are described in detail elsewhere (Office of Professional Standards Review, 1974; Goran et al., 1975). The quality of care assurance activities are described here in more detail.

To fulfill the quality assurance function, each hospital in a PSRO area must conduct four diagnosis-specific medical care evaluations (MCEs) annually. Selection of diagnoses can be made independently by each hospital, which could produce up to 28,000 different MCEs each year. In selecting diagnoses for study, physician committees in each hospital must consider the frequency of the diagnosis and the ability of medical care to alter the natural history of the diagnosed problem, and must have some reason to believe that the quality of care provided to patients with the selected diagnosis is currently inadequate. Again at the hospital's discretion, information used in the MCE may be drawn from either the medical record or patient interview; either process or outcome criteria may form the basis for the quality judgment. (Under federal sponsorship, the American Medical Association [AMA Criteria Development Project, 1976] coordinated the efforts of medical specialty organizations to develop criteria that local PSROs may use as guidelines. These are chiefly process criteria and rely on information available in the medical record.) Demonstrated deficiencies in care must be corrected and a second MCE is required to document improvements. Results of MCEs are reported to the area-wide PSRO, which then reports to the federal government.

In summary, the major features of the MCE are: 1) local determination by hospital physicians of the diagnostic categories for audit and of the criteria and standards against which care will be audited; 2) use of either process or outcome criteria against which to audit care (with a probable emphasis on process criteria); 3) reliance on
medical records or patient interviews, or both, for data on which to base judgments about quality; 4) review of quality of care based on groups or samples of patients, rather than a case-by-case approach; 5) correction of demonstrated deficiencies in care; and 6) reaudit to determine that deficiencies have been corrected.

Sample size will prove a major problem in using MCE results to improve quality of care. For the 59 conditional PSRO areas from which such data are available, the number of Medicaid enrollees ranges from about 5,400 to 504,000 and the number of Medicare eligibles ranges from about 2,200 to 385,000.* (No data are available on numbers of Maternal and Child Health eligibles.) Assuming that one in every ten eligibles/enrollees is admitted to the hospital annually, then the number of discharges across areas that would be subject to quality review would range from 760 to 89,000. If a diagnosis selected for review accounts for 2 percent of all discharges (a high estimate), the PSROs will have from 15 to 1760 cases in that diagnostic category. Given the wide range in the number of hospitals in PSRO areas, each hospital within the PSRO will have small numbers of cases to review for each of their MCEs. For most hospitals, consequently, it will be a statistical impossibility to document improvements in quality of care resulting from the diagnosis-specific MCEs.

The third PSRO function related to quality assurance is profile analysis, a retrospective review of patterns of care based on data from each hospital in the area aggregated on the basis of diagnosis, procedure, practitioner, and/or patients. This function is an adjunct to utilization review and MCE studies, insofar as it can help define problem areas in quality and assess the effectiveness of utilization review programs (Williams, 1976).

Summary

The PSRO program is a government-mandated cost-control and quality assurance system, organized and operated by physicians and hospitals at

a local level. Local control and local establishment of criteria and standards for peer review (as well as sample-size problems) will make it difficult, if not impossible, to compare findings of MCEs across PSRO areas. Results from the planned nationwide evaluation of the PSRO program, if it is performed, may be the only solution to this problem (Office of Research Evaluation and Planning, 1975). The evaluation design requires detailed examination of the care delivered to patients in 20 to 30 disease-specific categories. For each disease category, the effect of the PSRO system on cost of care, quality of care, and use of services will be determined. Because a standardized set of criteria and data collection instruments will be used, such an evaluation should eventually enable cross-regional comparisons. Thus far, however, the PSRO program has not been implemented in a manner that permits development of a strong evaluation design; the results of the program evaluation are therefore likely to be controversial.

Finally, the current scope of PSRO quality assurance activities is limited. MCEs review services delivered only to that portion of the population that is elderly, low-income, or both, and has been hospitalization. It will be hazardous, therefore, to use MCE findings to describe the quality of care received by the entire United States population.

**PERFORMANCE EVALUATION PROCEDURE**

The second major quality assurance program in the United States also focuses on hospital care; it was developed by a voluntary, non-governmental agency, the Joint Commission on Accreditation of Hospitals (JCAH). In the United States, hospitals are licensed by state governments, not by the federal government. Most of them also apply for JCAH accreditation, although application is strictly voluntary. The main incentives for accreditation are prestige and certification by the federal government for reimbursement of services delivered under Medicare and Medicaid. Until recently, JCAH accreditation was based largely on whether the hospital met certain structural criteria, such as posting a poison chart in the emergency room or following minimum staffing standards.
A few years ago, the JCAH devised a quality assurance program known as the Performance Evaluation Procedure (PEP) for Auditing and Improving Patient Care. Eventually, to keep its accreditation, each hospital will be required to conduct from four to 12 medical audits annually (the number of audits required is a function of the number of hospital admissions). These audits do not have to follow the PEP method, but must meet the six essential characteristics of an acceptable patient care evaluation procedure identified by the JCAH:

1) valid criteria that permit objective review of the care rendered to all patients must be established (i.e., they must be measurable and must be stated in terms of optimal achievable process or outcome, with an emphasis on expected patient outcome); 2) the practice of care is then audited in terms of the criteria; 3) the results of the audit are subjected to peer review, with recommendations for corrective actions if necessary; 4) corrective actions are implemented; 5) a follow-up process is instituted to make sure corrective actions are carried out; and 6) results are reported to the appropriate in-hospital committees and professionals.*

Currently, the JCAH is conducting a series of seminars--Trustees-Administrator-Physician Institutes (TAP Institutes, 1973) and Medical Audit Programs (MAP)--to describe the use of PEP, involve hospital personnel other than physicians in the program, and train appropriate personnel to carry out its functions. The method of performing a PEP has been described in great detail elsewhere (Jacobs and Jacobs, 1974), and will be discussed only briefly here.

Before doing a PEP, criteria and standards are established for a selected disease or procedure. Selection of the diagnostic or disease category for evaluation is done individually by each hospital. The criteria are applied on a case-by-case basis and must answer three questions: 1) Did the physician make the appropriate diagnostic and major intervention decisions? 2) How did the patient's actual discharge status (outcome) compare with that which was expected? 3) Do other clues (such as length of stay or complications during hospitali-

*Personal communication, Thomas Christoffel, Joint Commission on the Accreditation of Hospitals, December 1976.
zation) indicate that the patient received optimal care? Once the criteria and standards have been established, a trained auditor (employed by the hospital) reviews medical records to determine compliance with the standards; standards are developed to indicate those things that should always be done or never be done (either 100 percent or 0 percent compliance). When an audit indicates that standards have not been met, the record is sent to a physician review committee for reaudit. If the committee confirms deficiencies in care, follow-up efforts are undertaken to solve the problem.

In summary, the major features of PEP are: 1) local determination by hospital personnel of diagnostic categories for audit and of criteria and standards against which care will be audited; 2) emphasis on the outcomes of hospital care as well as on iatrogenic complications of care; 3) reliance on the medical record for data on which to base judgments about quality; 4) detailed examination of every case that fails to meet either standards set for management of the problem or standards set for end-results of hospital care; 5) reaudit to determine if deficiencies have been corrected; and 6) reporting of results to the JCAH rather than to a governmental regulatory body.

While the PSRO program and the JCAH are working on cooperative guidelines to prevent duplication of reporting requirements for MCEs and medical audits required by the JCAH, and to allow PEPs to qualify as MCEs,* there is now a difference in the approach taken by each program to quality assurance. The responsibility for PSRO review activities is delegated solely to physicians; the JCAH program encourages the assumption of responsibility by administrators and trustees as well as by physicians. The result is that PSRO review activities are likely to focus chiefly on physician care, and PEP audits are more likely to be patient-centered, reviewing all the care given to each patient. To prevent proliferation of quality assurance programs that examine only certain aspects of patient care, participation of all relevant providers in a single patient-centered audit

---

system is required. The process should also include patients, because decisions about the quality of medical care reflect value judgments that are clearly in the public domain. To date, no proposed quality assurance system has taken such an approach.

**HMO QUALITY ASSURANCE**

The second federally mandated quality assurance system was the result of the Health Maintenance Organization Act (HMO) of 1973 (P.L. 93-222, 1973). Briefly, an HMO is a prepaid group practice that delivers ambulatory and hospital care to an enrolled population. As a delivery system, the prepaid group practice competes with the fee-for-service system. To qualify for federal funding, HMOs must display certain characteristics defined in the 1973 legislation, which also made limited funds available for developmental activities.

Expressing concern that the prepaid group practice model might cut corners to stay within its budget by controlling delivery of expensive services or by fostering use of fewer services, the law requires each federally funded HMO to develop a quality assurance system. The system is to focus largely on outcomes of care for the defined population, and unlike those described above, is directed to review quality of both ambulatory and hospital services from the outset.

Because regulations for implementing HMO quality assurance systems were published only recently, there is no experience with their operation. However, a potential model for such systems exists in John Williamson's Health Accounting System (Williamson et al., 1975; Williamson, in preparation). Following Williamson's approach, local physicians first identify the health problem or problems that are associated with the greatest amount of preventable impairment in the relevant population; in applications of this approach to date, hypertension, urinary tract infection, and diabetes mellitus have accounted for almost 50 percent of the problems selected for quality assessment (Williamson, in preparation). Providers then establish outcome criteria against which to measure the extent of impairment reduced by medical care. (For hypertension, this standard might be
that 90 percent of patients with a diastolic pressure of > 105 should have a diastolic of ≤ 90 measured one year later.) Standards by which to judge the adequacy of treatment are then established, and outcomes are judged by patient follow-up for physiologic testing (in this example, a blood pressure reading), or completion of a questionnaire (in this example, to establish whether the patient has suffered a morbid event, such as a stroke). Follow-up data are then compared with the outcome standards. If quality is found wanting, the process of care (as documented in the medical record) is audited against process criteria for the problems in question that are developed by the physician group to determine the reasons for the poor outcome. At this stage, corrective actions are taken and the problem is reaudited to determine whether improvements have occurred. The practicality of implementing this system has recently been questioned, particularly for HMOs that serve low-income populations and those that lack the staff expertise to conduct the accounting procedure (Schroeder and Donaldson, 1976), and costs may prove too high when weighed against the benefits.

A rider to the HMO bill (P.L. 93-222, 1973) may prove even more important to quality assurance efforts than the requirement that HMOs install quality assurance systems. This rider authorizes the Secretary of Health, Education, and Welfare to:

- contract . . . for the conduct of a study to: 1) analyze past and present mechanisms . . . to assure the quality of health care; 2) provide a set of basic principles to be followed by any effective health care quality assurance system, including . . . specifications for the development of criteria and standards which relate to desired outcomes of care, and means for assessing the responsiveness of such care to the needs and perceptions of the consumers of such care; 3) provide an assessment of programs for improving the performance of health practitioners and institutions in providing high-quality health care, including a study of the effectiveness of sanctions and educational programs; 4) define the specific needs for a program of research and evaluation in health care quality assurance methods; and 5) provide methods for assessing the quality of health care from the point of view of consumers of such care.
Responsibility for implementing the study was assigned to the Institute of Medicine of the National Academy of Sciences, and the final report has recently been issued (Institute of Medicine, 1976). It is still too early to determine what effect the study will have on quality assurance in the United States.

CRISIS IN MALPRACTICE PREMIUMS

In recent years, premiums for physicians' malpractice insurance coverage have soared. In Ohio and other states, increases as high as 750 percent have been requested by insurors in a single year (Brook, Brutoco, and Williams, 1976). In southern California, as of January 1976, annual malpractice premium rates ranged from $4300 for pediatricians, to $7700 for general practitioners who do not perform surgery and $21,000 for those who do, to $36,000 for such "high-risk" specialties as neurosurgery and orthopedic surgery.

This rapid rise in premium rates provoked the "malpractice premium crisis" of the late 1970s. Physicians went on strike in several states and lobbied vigorously for legislation that would alleviate the financial burdens imposed by increased premiums. Some state legislatures have been reluctant to comply, because to some degree they view the threat of malpractice suits as a quality assurance mechanism in itself. (The relationship between malpractice and quality is unknown, but almost no physicians--less than 0.001 percent annually--are removed from the practice of medicine after losing a malpractice suit [Brook, Brutoco, and Williams, 1976]).

In some states, however, efforts to resolve the problem have included development of quality assurance activities in exchange for changes in the malpractice insurance system that would control premium increases. Some legislatures have passed bills that deal partially with the quality assurance aspects of the malpractice problem, requiring physicians, for example, to participate in continuing education programs in exchange for legislation to control premium costs. This approach may be effective if its underlying assumption is true—that inadequate knowledge produces poor-quality care. Unfortunately, there is little evidence to support such an assumption. For example,
Peterson et al. (1956), in their classic study of the quality of care provided by general practitioners, found no relationship between physicians' performance and their receiving and reading medical journals. Furthermore, the formal continuing education programs now in operation make little effort to determine the relevance of the topics they select to the problems faced by practicing physicians or their self-defined needs for additional knowledge (Brown and Uhl, 1970). Continuing education conducted in this vacuum may be both costly and ineffective. If the problems that prompt malpractice are problems of physician behavior and practice patterns, not lack of knowledge, a quality assurance mechanism that requires reaudit to demonstrate that documented deficiencies have been corrected will be more effective in promoting quality of care than will mandatory relicensure or continuing education. (For a discussion of the relationship between malpractice and quality of care, see Brook, Brutoco, and Williams, 1976).

CONCLUDING OBSERVATIONS

The following observations have clear implications for further research on quality assessment and assurance; they are discussed at greater length in Section IV.

1) Virtually all quality assurance activities review hospital care, not ambulatory care. 2) All quality assurance programs emphasize local selection of problems for review and local development of criteria and standards. 3) All review efforts concentrate on only a few disease conditions for quality assurance. 4) Virtually all programs propose to audit both process and outcome of care, but the emphasis (particularly in the PSRO program) is on process. 5) The medical record usually provides most of the data for quality assessment, as opposed to claims forms, patient follow-up interviews, and direct observation of the physician-patient encounter. 6) All quality assessment methods propose to use explicit criteria and standards against which to review care. 7) When the process of care is assessed, attention dwells on its technical aspects; the "art of care" is virtually ignored. 8) Data are sparse or nonexistent for evaluating the effectiveness of any of the quality assessment or assurance methods,
in terms of changes in provider behavior or in the health of the patients whose care is reviewed. 9) It is distinctly possible that duplication of expensive quality assurance activities will persist in the United States for the next several years.
IV. IMPLICATIONS FOR CONTINUING RESEARCH IN QUALITY ASSURANCE

HOSPITAL VERSUS AMBULATORY CARE REVIEW

Virtually all operational quality assurance programs emphasize the hospital component of medical care and ignore ambulatory care. This emphasis is a natural outgrowth of research in quality assessment, since most studies have been performed in hospital settings. The validity of this approach can be questioned. Will quality assurance activities conducted in the hospital increase the health of people more than would similar activities in the ambulatory component of medical care? The answer to this question is unknown, but intuitively one might surmise that the answer is yes, on the grounds that the drugs and procedures used in the hospital are more dangerous (their safe therapeutic range is narrower and side effects are more severe).

Recent studies by Payne et al. (1976) dispute this intuitive judgment. Their studies assessed the quality of care provided by physicians in Hawaii to patients in both hospital and office-based settings. They measured the quality of care rendered to patients in 20 diagnostic categories by comparing data in the medical record with a list of disease-specific process criteria generated by members of the Hawaii Medical Society (see Table 1 for an example). They weighted and combined these criteria into an overall physician performance index. The index was scaled so that compliance with all criteria would produce a score of 100, and compliance with none a score of zero. Across all diagnoses, the average performance index was 77 for the hospital component of care and 44 for office-based practice. Hospital care was thus judged to be closer to the "ideal" than was ambulatory care. A literature review of several hundred research studies of health services supports this general conclusion (Brook, 1973b).

It should be noted, however, that physician performance in hospital versus ambulatory care, as defined by Payne's scale, may not correlate linearly with benefits in health status. Consequently, the
Table 1

PROCESS CRITERIA FOR PNEUMONIA:
HAWAII MEDICAL ASSOCIATION

Indications for Admission: Presence of proven or suspected pneumonia

Services Recommended:

1. History: Specific reference to:
   a. Character of sputum .................................. 2.0
   b. Pain in chest ....................................... 2.0
   c. Duration and degree of fever .................. 2.0
   d. Onset of illness .................................. 1.0
   e. Previous episodes and social history .......... 1.0
   f. Contact history ................................... 1.0

2. Physical Examination: Specific reference to:
   a. Breath sounds, character, presence or absence of... 2.0
   b. Friction rub ....................................... 1.0
   c. Rales .............................................. 1.0
   d. Chest movements ................................. 1.0
   e. Percussion ....................................... 1.0
   f. Cyanosis ......................................... 2.0
   g. Vital signs—temperature, pulse, respiration,
      blood pressure .................................... 2.0
   h. Character of respiration ........................ 2.0

3. Laboratory:
   a. C.B.C. ............................................. 2.0
   b. Blood culture in seriously ill patient (temperature
      of over 104°, cyanotic, needs oxygen) .......... 3.0
   c. Sputum or throat culture with sensitivities .. 3.0

4. Roentgenology:
   a. PA and lateral of chest on admission .......... 3.0
   b. Follow-up chest X-ray in 7 days or
      before discharge .................................. 2.0

Therapy: Appropriate antibiotics (not sulfa drug or
chloramphenicol) ....................................... 3.0

SOURCE: Adapted from Payne et al., 1976.
results may not suggest, as one might gather from a first reading, that ambulatory care stands in greater need of improvement than does hospital care. Studies are needed that estimate the incremental improvement in health obtained for the same marginal investment in hospital care and in ambulatory care.

The decision to assess only hospital care in the PSRO program was probably based on practical considerations. It is far easier to perform an audit in a hospital than in the office of a solo practitioner, and costs of hospital care represent the greatest proportion of Medicare and Medicaid budgets. Nevertheless, this decision may partially invalidate the audit judgments. For instance, suppose a PSRO established process criteria for the treatment of patients with pneumonia. These criteria certainly would require performance of a sputum culture and a chest x-ray. If two doctors complied with these criteria, one in his office before the patient was admitted to the hospital and the other after the patient was admitted, and if only data in the hospital record were used in judging the quality of care, it would appear that these physicians had provided different levels of care. Such erroneous conclusions could lead to practices that would result in higher (i.e., worse) cost/effectiveness ratios for the medical care system. The physician who correctly performed these tests on an outpatient basis may feel compelled, because of the quality assurance system, to repeat them on an inpatient basis. This course of action may be the least expensive in terms of the physician's time, since reviewing the case with the hospital audit committee would be more time-consuming than repeating the tests. The potential duplication of services and consequent increase in costs as the result of a hospital-based quality assurance system need to be confirmed or denied by proper research. Eventually, quality of care assessment must integrate the review of both ambulatory and hospital care.

LOCAL CONTROL

The quality assurance programs now operating in the United States emphasize local selection of problems for review and local establish-
ment of criteria and standards. Three reasons are usually given for this emphasis: 1) local physicians are more familiar with the medical care needs and problems of the populations they serve; 2) physicians display a built-in antipathy toward use of nationally mandated or "federal" criteria; and 3) even though they may use the national criteria, they would be less likely to "internalize" them and would be more likely to change their practice behaviors to meet locally established criteria.

There is no evidence to support such reasoning. If problems in the delivery of medical care are much the same everywhere, complete local freedom of selection may prove to be unnecessary and time-consuming. As a case in point, every community study of the adequacy of blood pressure control for hypertensive patients has found gross deficiencies in care (Brook, 1973b). Furthermore, because information on socioeconomic and environmental conditions in a medical care area is readily available, "outsiders" should be able to identify important medical care problems. Local physicians may lack the epidemiologic experience necessary to use this information in an equally valid manner.

There is no evidence that physician behavior is more responsive to local than national standards, or that the reverse is true. Because this is such a crucial point, it should be tested rigorously through controlled experimentation. Furthermore, establishment of local criteria and standards raises ethical and moral questions. Should a rural physician who prescribes chloramphenicol for a cold be permitted to practice as usual, while an urban physician would suffer harsh sanctions for doing the same? Criteria will vary according to local social and economic conditions; for instance, a patient with pneumonia who has an unheated home may require a longer hospital stay than the patient who can be discharged to a home with central heating. Is it justifiable, however, to reject valid criteria and standards for the delivery of good quality of care based only on the fact that local physicians do not accept such criteria as indicative of good care? Sooner or later, ethics may require that those process criteria shown to be related to health care are uniformly applied throughout the country;
likewise, those criteria that do not meet this level of validity should not be used to judge care, regardless of the opinion of local physicians.

SCOPE OF REVIEW ACTIVITIES

All operational quality assurance systems select only a few conditions or procedures for intensive study. The vast majority of conditions or procedures go unstudied. Nonetheless, there is a strong tendency to generalize findings based on studies of the care delivered to patients in one or two diagnostic categories to all the care given throughout an institution or region. The validity of such a generalization must be tested rigorously. Only a few research studies have attempted to review the quality of care provided by individual physicians to groups of patients with selected diagnoses, and then determine whether the results were generalizable to the care the same physicians gave to patients with other conditions. When such studies have been performed, such as those by Brook (1973a) and Payne et al. (1976), they have demonstrated that the level of care can vary substantially, even between groups of patients with similar conditions (in terms of chronicity) who are treated by the same set of providers. For instance, in Brook's study, the medical house staff provided a significantly lower level of care to patients with urinary tract infection than they did to patients with ulcerated lesions in the stomach or duodenum. Yet both of these conditions are common chronic medical problems and are often chosen as indicator conditions for quality review.

TYPE AND SOURCE OF DATA

Several quality assurance systems propose to use some combination of process and outcome data to assess quality of care, and to abstract those data from the medical record. Most systems, PSRO in particular, emphasize the use of process data. The question that must be addressed is how the validity of the results of the quality assessment varies as a function both of the type of data used and the source from which they are obtained. For instance, the process of ambulatory care has
been assessed on the basis of data in insurance claims forms, but can also be measured by examining the medical record. How closely the assessments based on different data sources correspond is unknown.

The findings of a recent study underscore the importance of these issues. The Joint Committee on Quality Assurance of Ambulatory Health Care for Children and Youth (Osborne and Thompson, 1975) reviewed the care given by board-certified pediatricians to children with one of six conditions, such as well-child examination and asthma. The level of quality of care was ascertained by comparing the practitioners' medical records with explicit process criteria agreed to by the members of this Committee. The study found that the care given by supposedly excellent practitioners was seriously inadequate; however, after discussing the recording of process items with the physicians, the Committee concluded that the recording procedures were deficient, not the care. No study was conducted to test this conclusion by comparing compliance with criteria as determined by observing physician practice with compliance as determined by examining the medical record.

An answer to that question is vital; otherwise, it will be too easy to dismiss findings of deficient care based on data obtained from the medical record. Lyons and Payne (1974) partially answered this question by demonstrating, in a record review, that physicians who complied with criteria that required both doing and recording (such as examining the fundi in a hypertensive patient) also had higher scores on those criteria that required physician action but were recorded by office personnel (such as ordering measurement of the potassium level in a hypertensive patient). The correlations were only weakly positive, however, and varied by physician specialty; they suggest that the relationship between providing high-quality medical care and keeping good medical records may not be strong.

Unless the recording of the process of care is upgraded, it is not desirable to press strongly for assessment of the quality of ambulatory care (i.e., beyond examination by extremely simple criteria of whether appropriate laboratory tests were done or correct therapy was given). Whether such recording is beneficial and should be paid
for by increasing the cost of an ambulatory visit is also debatable. For instance, if the patient presents with a cut on the hand and the physician tests for sensation, mobility, and strength but does not record this information, should he be penalized by a quality assurance system that derives its data from the medical record and mandates testing for these attributes before suturing the lesion? Such a quality assurance system may do more harm than good if it compels the physician to spend less time with the patient and more time with the record.

Despite the common focus of quality assurance systems on process assessment, the issue of which type of data—structure, process, or outcome—produces the most valid result has not yet been resolved. The results of the quality assessment clearly depend on the type of data used. Fessel and Van Brunt (1972) studied the quality of care given to patients who exhibited symptoms of appendicitis and were treated in one of three hospitals. They judged the process of care given to these patients by comparing the level of care as recorded in the medical record with an explicit set of process criteria. They then judged the outcome of care by examining the proportion of patients operated on who actually had acute appendicitis (Table 2). The hospital scoring lowest on the process assessment (Hospital C) had the best score when judged by the outcome assessment. In this case, treatment at the hospital that performed fewer inappropriate operations is preferable to the hospital that had more complete medical records.

Brook (1973b) studied the quality of care given to 296 patients who had either hypertension, a urinary tract infection, or an ulcerated lesion in the stomach or duodenum. The quality of care given to these patients was assessed independently by five different methods, relying on physician judges: 1) process information judged by implicit criteria; 2) outcome information judged by implicit criteria; 3) process and outcome information judged by implicit criteria; 4) process information judged by explicit criteria; and 5) outcome information judged by explicit criteria (see Table 3). Implicit criteria refer to the gestalt, unenunciated opinions of the reviewing physi-
Table 2

PROCESS AND OUTCOME ASSESSMENTS OF THE QUALITY OF CARE GIVEN TO PATIENTS WITH ACUTE APPENDICITIS WHO WERE TREATED AT ONE OF THREE HOSPITALS

Part A
Recorded Data on the Process of Care

<table>
<thead>
<tr>
<th>Hospital</th>
<th>Probability difference due to chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Referred rebound</td>
<td>82</td>
</tr>
<tr>
<td>Guarding</td>
<td>96</td>
</tr>
<tr>
<td>Psoas sign</td>
<td>48</td>
</tr>
<tr>
<td>Obturator sign</td>
<td>32</td>
</tr>
<tr>
<td>Bowel sounds</td>
<td>98</td>
</tr>
<tr>
<td>Organ enlargement</td>
<td>30</td>
</tr>
<tr>
<td>Rectum</td>
<td></td>
</tr>
<tr>
<td>Tender on right</td>
<td>68</td>
</tr>
<tr>
<td>Other tenderness</td>
<td>12</td>
</tr>
<tr>
<td>Normal</td>
<td>20</td>
</tr>
</tbody>
</table>

Part B
Diagnostic Outcomes

<table>
<thead>
<tr>
<th>Hospital</th>
<th>Number of patients</th>
<th>Pathologically proved acute appendicitis %</th>
<th>No pathological findings %</th>
<th>Diagnosis other than appendicitis %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>466</td>
<td>82.2</td>
<td>13.6</td>
<td>4.1</td>
</tr>
<tr>
<td>B</td>
<td>167</td>
<td>83.1</td>
<td>13.7</td>
<td>4.2</td>
</tr>
<tr>
<td>C</td>
<td>104</td>
<td>89.4</td>
<td>10.6</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3

SUMMARY OF THE ASSESSMENT OF QUALITY OF CARE AS MEASURED BY EACH OF THE FIVE METHODS STUDIED FOR 296 PATIENTS WITH EITHER HYPERTENSION, URINARY TRACT INFECTION, OR ULCERATED LESION OF THE STOMACH OR DUODENUM

<table>
<thead>
<tr>
<th>Quality Assessment Method</th>
<th>Number of Patients</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Implicit process judgment</td>
<td>69</td>
<td>23</td>
</tr>
<tr>
<td>II. Implicit outcome judgment</td>
<td>187</td>
<td>63</td>
</tr>
<tr>
<td>III. Implicit process and outcome judgment</td>
<td>80</td>
<td>27</td>
</tr>
<tr>
<td>IV. Explicit process judgment</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>V. Explicit outcome judgment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. for urinary infection</td>
<td>42</td>
<td>40</td>
</tr>
<tr>
<td>b. for hypertension</td>
<td>50</td>
<td>44</td>
</tr>
</tbody>
</table>

SOURCE: Modified from Brook, 1973b.
cian(s) as to what care of optimal quality can be expected to achieve in terms of process, outcome, or both; explicit criteria are those agreed to a priori by reviewing physicians and written down for comparison with information in the record. The judgment of the level of quality of care provided varied considerably as a function of the methods used. Fewer than 2 percent of the patients were judged to have received adequate care by the explicit process method, while the implicit outcome method indicated that 63 percent had received adequate care.

The American Society of Internal Medicine (Hare and Barnoon, 1973) completed a study in which practicing internists in six areas of the United States established process criteria for six common conditions. Subsequently, the Society assessed the care given to patients with these conditions by those practitioners who had established the criteria. An inverse correlation was found between the weight given to a criterion and compliance with it. Apparently, the physicians did not practice what they preached. The explanation may be either that they provided care of poor quality, or that the method of assessment was unreliable.

The results of the work by the Joint Committee, Brook, Fessel and Van Brunt, and the American Society of Internal Medicine illustrate how strongly the findings of a quality of care assessment depend on the method used. First, large discrepancies will be found as a function of whether process or outcome information is used in the assessment. It is not clear which type of information produces the most valid judgment of care, but it is clear that process assessment produces much harsher judgment. The reason is obvious when one considers the probabilistic nature of medicine. Process criteria as currently developed rarely take into account the possibility, given the natural history of disease, that a "paradoxically good" outcome may occur even if the process of care is considered inadequate. If patients with bacterial pneumonia were incorrectly diagnosed and not treated with antibiotics, most of them would recover. An outcome assessment focused on recovery would indicate that these patients received good care; a process assessment would indicate care was poor.
The "truth" in this case would be more on the side of the process assessment, although some of the process criteria the physicians selected may have been of questionable validity. For instance, is it medically necessary that every patient with pneumonia receive all the treatment procedures indicated by the process criteria listed in Table 1? If lists such as this typical example were used to assess quality of care, the number of "things" to be done to patients would increase two- to threefold for ambulatory patients and by a factor of one and one-half for hospital patients. This increased activity would probably raise the percentage of the GNP spent on health to over 15 percent, yet would improve health very little, since the quality of care assessment included invalid process criteria.

Why do practicing physicians accept, on paper, invalid process criteria? The American Society of Internal Medicine's study certainly suggests that physicians do not follow these criteria in their daily practice. Perhaps the reason is the "academic" approach of the techniques used to produce these criteria. This issue, too, calls for more research.

Finally, the artificiality of the weighting scheme used to derive an overall quality score represents a serious deficiency in method. If all the criteria in Table 1 were met except that of giving an antibiotic, the physician could achieve a near-perfect score and yet the patient might die. Obviously, the weighting scheme does not replicate the clinician's decision-making process. The artificiality of this weighting scheme may explain why implicit judgments of the process of care, whereby the care is examined in a gestalt manner, produce results different from and probably more valid than those of explicit review (Table 3). At least the implicit process judgment is closer to the physician's decision-making process than is the explicit process judgment. In an attempt to advance the state of the art, the UCLA EMCRO (Greenfield et al., 1975) has used several decision-making principles in developing process criteria. Their explicit process criteria are branched so that an action indicated by one criterion often depends on the results of a previous action. In many cases, a physician can use any one of the multiple pathways in reaching a deci-
sion during the process of care. The feasibility of applying this complex system to the average medical record needs further testing.

TECHNIQUE VERSUS ART OF MEDICAL CARE

All of the approaches described above concentrate on measuring the technical aspect of the process of care; recent advances in measuring the art of care from the patient's perspective by Hulka and others (for example, see Zyzanski, Hulka, and Cassel, 1974) and by Ware and others (Ware and Snyder, 1975; Ware, Snyder, and Wright, 1976a, 1976b) have not been taken advantage of in developing quality assessment methods. As opposed to the technical aspect of process, which generally refers to the performance of preventive, diagnostic, and therapeutic measures related to the patient's symptoms or problems, the art of care refers to the manner of physician care relative to the patient as an individual, as measured by the physician's sensitivity, openness, and nonauthoritarian manner. Optimal quality in the art of care might be expected to positively influence: 1) patient compliance; 2) willingness to discuss sensitive problems; 3) willingness to change health behaviors associated with increased morbidity and mortality; and 4) appropriate and timely use of medical services. As measures of the art of care from the patient's perspective are shown to be reliable and valid, they should be incorporated into operational quality assurance systems so that at least some emphasis is placed on improving the art as well as the technical component of care.

EVALUATION OF QUALITY ASSURANCE SYSTEMS

Unfortunately, virtually all systems for assessing and assuring quality have been implemented in a manner that will make it difficult to evaluate their impact on quality and on health status. Consequently, the effectiveness of these systems in changing physician behavior or improving people's health will remain largely unknown. Only a few studies have documented any changes at all (Brook and Williams, 1976), and others question whether the documented changes are temporary or permanent. Furthermore, little is known about the
most effective way to change either system or provider behavior to produce higher levels of quality. How should quality assessment information be reported? Should the assessment be done on a group or individual basis? Should data be aggregated by specialty, by region, by hospital, or by community? What should be done about "bad doctors"? Should they be removed from the practice of medicine or should continuing education be mandated? Efficient and effective mechanisms for changing unnecessary or inadequate provider behavior must be found. Quality assurance nationwide is likely to consume 2 to 5 percent of available medical care dollars. If change could be accomplished by data aggregated at the community level rather than disaggregated at the individual provider level, sampling could be used and operational expenses decreased. Research along these lines is clearly necessary.

DUPLICATION OF SYSTEMS

Finally, it is apparent that duplicate or overlapping quality assurance systems are being established in the United States. Because of economic considerations, federal regulatory systems will likely come to the fore, and the role of voluntary agencies such as the JCAH may disappear. If and when national health insurance becomes a reality, the PSRO program will likely expand to review the care provided to the entire population. As the PSRO review activities expand to include ambulatory care, as eventually they must, they may duplicate those activities mandated by the HMO legislation. The decision on how best to eliminate this costly duplication of activities will be partly political, but it should also be made at least partially on the basis of information about the cost-effectiveness of the different approaches and their effect on quality of care and the health of the population.
V. CONCLUSIONS

The United States is currently implementing a series of quality assurance programs. Their short-term effect on quality of care and the health of the American public will depend on rigorous program evaluation and feedback of results that will enable improvements and correction of deficiencies. Without such feedback, the programs may prove costly yet produce no demonstrable improvements in health. Their long-term influence will depend on development of more valid and reliable methods of assessing quality of care and of more efficient and effective ways to modify provider and patient behaviors.

The knowledge for accomplishing these ends will come only with completion of careful research. The search for new and better methods must be interdisciplinary, and should not be confined to a single system of medical care delivery. The issue of measuring and improving quality of care is one of continuing international importance, regardless of whether care is provided on a fee-for-service basis, is partially subsidized, or is nationally organized.

Although further research is without a doubt necessary, this report should not be interpreted as advocating postponement of quality assurance programs until they can be rendered virtually flawless. Present-day quality assessment methods are able to detect major deficiencies in quality of care, as numerous studies done here and abroad will attest. Until these gross deficiencies are identified and corrected, more sophisticated techniques may not be necessary.

The most serious threat to the positive effect of quality assurance programs on care and health status lies with failure to use the results of quality assessment studies to improve care. Many studies have described deficiencies in care, but little has been done with the information. This can no longer be allowed. Effective quality assurance programs must take full advantage of feedback mechanisms, such as the reaudit procedures required by PSRO, PEP, and Williamson's Health Accounting System, to ensure quality and improve health status.
REFERENCES


National Center for Health Services Research and Development, and
Arthur D. Little, Inc. EMCRP Program. DHEW Publication No. HSM
73-3017. Rockville, Maryland, Department of Health, Education, and

Nightingale, F. Mortality of the British Army at home and abroad and
during the Russian War as compared with the mortality of the civil
Appointed to Inquire into the Regulations Affecting the Sanitary

Nightingale, F. Proposal for Improved Statistics of Surgical Opera-

Rockville, Maryland, Department of Health, Education, and Welfare,
1974.

Office of Research Evaluation and Planning, Office of Professional
Standards Review. PSRO Program Evaluation Plan. Rockville,

Osborne, C. E., and H. C. Thompson. Criteria for evaluation of ambu-
latory child health care by chart audit: Development and testing of

P.L. 92-603. 92d Congress of the United States of America, H.R. 1,
Office, pp. 101-114.

the 93d Congress, 1st sess. Washington, D.C., U.S. Government
Printing Office, pp. 1015-1042.

Payne, B. C., et al. The Quality of Medical Care: Evaluation and
Development. Chicago, Hospital Research and Educational Trust,
1976.

1956.


Sanazaro, P. J., et al. Research and development in quality assur-
ance: The Experimental Medical Care Review Organization program.

Sasuly, R., and C. E. Hopkins. A medical society-sponsored compre-


Williamson, J. W. Medical Care Outcomes: Assessment and Improvement. (In preparation).