DRAFT - Physician Payment Review Commission - September 28,1988
Background Paper for Conference

IMPROVING THE QUALITY OF CARE: CLINICAL RESEARCH AND PRACTICE GUIDELINES

Congress will soon consider proposals to reform Medicare physician payment or, more generally, the fee-for-service sector of Medicare. It created the Physician Payment Review Commission to give it advice on how this might be done. The task can be conceived as having two major parts or objectives: change the way charges or fees are set, and implement a strategy to gain some control over the rate of growth of expenditures for physician services without compromising access and quality of care for beneficiaries.

The Commission's work is proceeding along two tracks. First, we are developing a proposal for a fee schedule. A fee schedule would be less complex and easier to administer than the present system, and it would put fees on a rational basis. It would reduce the distortions in financial incentives for how physicians practice, where they locate, and what specialty they select.

However, a fee schedule alone will not control the costs of physician services, nor insure that patients receive the services most appropriate for their needs. The rapid increases in expenditures per capita over the past decade have been driven more by increases in the volume of services than by price increases. Though a fee schedule would reduce the effects of distorted financial incentives on physicians' decisions, more direct measures are also needed to improve the appropriateness of the services that patients receive.

Consequently, the Commission is also developing a strategy to improve the quality of care by programs to increase the appropriateness of services provided to Medicare beneficiaries. The Conference on Practice Guidelines is part of our work to evaluate the potential of an important element of such an approach: to increase through clinical research our knowledge of what services are effective for patients and, through practice guidelines, to make this information available to physicians, patients, and those who review utilization and quality of care. Though this would not be accomplished easily, it appears to offer considerable potential to increase the quality of care, and reducing inappropriate and unnecessary services would reduce costs.

This paper will provide a rationale for increasing the resources devoted to clinical research and evaluation of medical practices, and for the development and use of practice guidelines. It will briefly review the evidence that patients now receive some services that are unnecessary or inappropriate for them, and

explain how quality of care could be improved by giving physicians and others better information on appropriate care through clinical research and practice guidelines. It will also discuss several arguments against this approach. Finally, it will discuss potential roles that the federal government might play in funding additional clinical research and development of practice guidelines.

The paper has four appendixes that provide additional detail. Reference will also be made to several discussion papers prepared for the Conference:

> "Appropriateness of Acute Medical Care for the Elderly" by Dr. Robert Brook et al,
>
> "Practice Guidelines and Practicing Medicine: Are They Compatible" by Dr. Robert Brook,
>
> "Leadership in the Development of Practice Guidelines: The Role of the Federal Government and Others" by Mr. Lawrence Lewin and Ms. Jane Erickson, and
>
> "Methods for Designing Guidelines" by Dr. David Eddy.

APPROPRIATENESS, QUALITY, AND COSTS OF CARE

There is increasing evidence that beneficiaries receive some services that are unnecessary and fail to receive some services that would benefit them. Over many years, a substantial research literature has documented unnecessary laboratory tests, radiological procedures, hospital admissions and days of care, surgical procedures, and drugs (Chassin et al, 1986b; Eisenberg, 1986; Schroeder, 1987). More recent work has focused on several specific procedures, including carotid endarterectomy (Winslow, 1988), upper gastrointestinal endoscopy (Chassin et al, 1987b), insertion of cardiac pacemakers (Greenspan et al, 1988), cardiac catheterization (Chassin et al, 1987a), and coronary artery bypass operations (Winslow et al, 1988a; Graboys et al, 1988). These studies have shown that about 10 - 30% of recipients were unlikely to benefit because the procedures were performed for inappropriate reasons. An additional fraction were performed for indications that experts could not agree were clearly appropriate or inappropriate.

> The accompanying paper "Appropriateness of Acute Medical Care for the Elderly" by Dr. Brook and colleagues reviews this evidence in detail for the elderly population.

These inappropriate procedures reduce the quality of care and increase cost. Inappropriate invasive procedures may involve

substantial risks to the patient without accompanying benefit. For example, the study of carotid endarterectomy found that about ten percent of patients undergoing the procedure suffered perioperative strokes or died.

The cost of these unnecessary services is also substantial. For instance, Medicare spends more than $300 million per year on physicians' charges for coronary artery bypass operations. If just ten percent of these are inappropriate, $30 million in physicians' charges and millions more for the associated hospital costs could be saved or spent for more beneficial services.

Substantial variation in the use of services among geographic areas has been documented for hospitalization and for many procedures. For instance, Wennberg has shown per capita variations as great as six-fold among communities in New England in the use of procedures like hysterectomy and prostatectomy (Wennberg et al, 1982; Wennberg, 1984). Chassin and colleagues demonstrated similar variations in the use of many procedures for larger areas (Chassin et al, 1986b). This variation is cited as additional evidence that patients are receiving unnecessary services in some areas, or failing to receive needed services in others, or both. They also demonstrate that there are large differences in expenditures for a procedure among areas generally comparable in medical needs.

Why do physicians and patients make decisions that are judged to be inappropriate, that increase costs without benefit and may even harm the patient? There are a number of factors. Clinical decisions to use services for patients are influenced by the organization of practice and the way the physician is paid, by the availability of hospital beds and physicians, and other factors. This paper will focus on one factor that is important whether care is provided through HMOs or fee for services, whether in rural areas without a hospital or a tertiary medical center with the latest in technology: the information available to physicians, patients, and others about the risks and benefits of medical services and procedures. It has become apparent that the physician and patient often simply do not have the right information to confidently predict whether a particular service or procedure can be expected to benefit the patient.

In some cases, even the experts do not know whether a procedure is likely to be effective -- not enough information is available because the clinical research to fully determine the risks and benefits of the procedure has not been done. In others, the information is available in the research literature, but many individual practicing physicians may not have that information, or may not use it.

Scholars have attributed much of the geographic variation and the use of unnecessary services to uncertainty among physicians as to

what services are appropriate for their patients (Wennberg et al, 1982, 1985, 1987; Eddy, 1984; Eisenberg, 1986). Wennberg has noted that variation is lower for services for which physicians have a broad consensus about appropriate use, such as the need to hospitalize patients with hip fracture and suspected myocardial infarction. Uncertainty may also play a substantial role in use of services for defensive medicine (Eisenberg, 1986).

Incomplete knowledge of the risks and benefits of medical services and procedures also limits the effectiveness of utilization and quality review. For instance, if PROs using sound criteria inform a physician that a particular surgical procedure is not indicated for a patient, or deny payment for that inappropriate procedure, the physician will learn from the experience. Sound review will reinforce good practice.

However, review criteria are not always based on sound evidence, so much review is open to the criticism that it is arbitrary and inhibits good quality care. Most utilization review is now based on statistical norms or averages. It sanctifies these statistical norms as though they were clinical "standards" without knowing what effects these practices have on patient outcomes. Review of utilization and quality should be based on knowledge of appropriate medical practice. Since reviewers influence physicians' decisions and thus both the cost and quality of care, they should base the review on the best medical knowledge available.

Clinical research and practice guidelines

There are two ways to increase what we know about the risks and benefits of medical services: clinical research on effectiveness, and development of practice guidelines that draw on this research and on the clinical experience of practicing physicians.

First, substantial additional work to evaluate the safety, efficacy, and effectiveness of clinical practices is clearly needed to support efforts to improve medical care. Current funding for these activities is small, especially when compared to expenditures for the Medicare program. As the Institute of Medicine has noted (1985), funding for clinical research is meager, and funding for evaluation of medical practices is only a fraction of that.

Second, we must make better use of the results of clinical research and of the knowledge that is gained through clinical experience. Neither the research literature nor the collective experience of practicing physicians are easily accessible to the individual practicing physician. This information on effectiveness of medical services and procedures can be made available to the clinician through its translation into practice

guidelines.[1]

Practice guidelines are not new. They have long been used to provide clinical information to physicians. Examples include protocols for diagnosis and management of cancer and recommendations for the use of preventive services like screening mammography. But there has been a recent increase in development of practice guidelines for a variety of medical and surgical procedures. Guidelines recently developed by a joint project of the American College of Cardiology and the American Heart Association for the use of coronary angiography provide an example (ACC/AHA Task Force on Assessment of Cardiovascular Procedures, 1987). These guidelines divide possible indications for coronary angiography into three categories:

- conditions for which there is general agreement that coronary angiography is justified,
- conditions for which coronary angiography is frequently performed, but there is a divergence of opinion with respect to its justification in terms of value and appropriateness, and
- conditions for which there is general agreement that coronary angiography is not ordinarily justified.

Guidelines like these could be used by a physician to help decide when to recommend coronary angiography to patients. If they were to be used as the basis for utilization and quality review, a reviewer might deny approval for an indication that was in the third category pending a more complete investigation.[2]

Appendix 1 explains in greater detail what practice guidelines are and gives two examples.

---

[1] Some refer to formally developed practice guidelines as "clinical standards." The former term is used here to note that recommendations for proper clinical care embody varying degrees of uncertainty. Those in which there is great confidence could be offered as standards to which all should hold, but most should remain as guidelines until their validity as predictors of patient outcomes has been established.

[2] Many in the medical community advocate that guidelines initially be used to inform patients and physicians. Only after they had been accepted by the medical community as valid would they be used for review and determination of payment.

DEVELOPMENT OF PRACTICE GUIDELINES AND THE ROLE OF THE FEDERAL GOVERNMENT

For additional guidelines to be developed, a method must be chosen and a sponsor must take responsibility for the task. There are several potential roles for the federal government in this process: to fund the clinical research upon which guidelines would be based, to sponsor the research needed to develop and evaluate methods for constructing guidelines, to develop the guidelines, and to provide funds for these activities.

> Appendix 2 suggests desirable characteristics of guidelines and of methods to develop them. The accompanying paper by Dr. Eddy discusses methods in greater detail. Appendix 3 suggests what characteristics suitable sponsors would have. The accompanying paper by Mr. Lewin and Ms. Erickson discusses sponsorship or leadership in the development of guidelines.

Development of Guidelines

The soundness of guidelines and their acceptability to physicians will depend on the method used to develop them and the credibility and authority of the sponsor. The method should be a rigorous, structured process that synthesizes the information in the scientific literature, extends it through the knowledge of expert physicians, and expresses the information in specific, precise, comprehensive guidelines. The method must minimize effects of possible biases of the participants.

Several existing methods have many of the characteristics desired, including those used by the American College of Physicians, the American College of Cardiology and American Heart Association, and the one developed by the RAND investigators for the Health Services Utilization project (Brook et al, 1986). However, each of these methods could be improved and refined.

A sponsor for development of guidelines must establish an administrative structure to carry out the project, secure the necessary funds and resources including physicians and other professionals, choose or develop a reliable and valid method, and oversee its proper use. The sponsoring organization must be authoritative and respected for having high ethical and scientific standards, objective and reasonably free of self-interest, either professional or financial.

Most formal guidelines developed to date have come from specialty societies. Several specialty societies including the American College of Physicians and the American College of Cardiology have been developing guidelines for several years, and other specialty societies have recently begun projects to develop them. Specialty societies have the physician expertise and credibility

with the physician community to do the job. However, the potential for or appearance of bias from a particular society's vested interests must be countered by including a range of specialties in the process, using a sound method designed to minimize bias, and by the presence of oversight by a respected sponsor. Specialty societies, either alone or in a consortium, may not have the resources to sponsor a large enough effort to have a substantial effect on the practice of medicine.

Other organizations have contributed to the development of guidelines, particularly the development of sound methods. The RAND Corporation developed a method and applied it by developing guidelines for several medical and surgical procedures. A consortium of academic health centers is planning to evaluate and refine available methods and to apply the result in developing a small number of additional guidelines. While these projects demonstrate the research, development, and evaluation role that organizations like these can play, they are less likely to mount a large scale project to produce guidelines for many services and procedures.

The federal government has not developed formal guidelines for medical practice. The National Institutes of Health has held more than 60 Consensus Development Conferences to assess the safety, efficacy, and effectiveness of specific procedures and practices (Perry, 1987; Mullen and Jacoby, 1987). The consensus statements issued following these conferences are assessments of the safety, efficacy, and effectiveness of the technology rather than formal practice guidelines, though they are intended to influence medical practice. It has been suggested that changes in the Consensus Development Conferences should be made to increase the specificity of the recommendations (Mullen and Jacoby, 1987).

Role of the Federal Government

The federal government has an important role to play in laying the groundwork for higher quality and more efficient care, both through clinical research and the development of practice guidelines. It is a payer and the overseer of care for Medicare beneficiaries and Medicaid recipients. More importantly, it shares responsibility for the health and welfare of all of the American people.

> Mr. Lewin and Ms. Erickson's discussion paper discusses the role of the federal government in greater depth.

It seems clearly appropriate for the federal government to support additional clinical research to develop our knowledge of the effectiveness of medical practices. The National Institutes of Health (NIH) and the National Center for Health Services Research and Health Care Technology Assessment (NCHSR) now house

or administer funding for much of the nation's health related research. Expanded research to evaluate clinical practices could be conducted and funded through these agencies and the Office of Research and Demonstrations of the Health Care Financing Administration (HCFA). HCFA has recently announced an Effectiveness Initiative to expand its funding of clinical research. The work needed to evaluate and refine the methods to be used to develop practice guidelines--the next step in the process--could also be conducted through these agencies.

Federal funding would also be required if there is to be a substantial effort to develop the guidelines themselves, but it is not clear what the precise role of the federal government should be. The government could mobilize the resources to develop guidelines or it could contract with others for their development. There is a great deal of expertise available, particularly through the NIH. HCFA has access to a wealth of data generated from clinical practice that should prove useful.

However, in contrast to the work to evaluate and refine methods, there are disadvantages to internal governmental development or direct contracting for development of guidelines. The physician community views the government, particularly HCFA, as primarily interested in guidelines as a means of reducing costs. The medical community would be skeptical of guidelines developed through HCFA, and might be reluctant to cooperate in their development.

Alternatively, the federal government could support and build upon work already begun by the medical profession. A number of individual specialty societies are developing or planning to develop guidelines, and medical associations that cross specialty lines, like the American Medical Association and the Council of Medical Specialty Societies, are considering projects to develop guidelines. These efforts would benefit from oversight and guidance to ensure that the methods used were of high quality and the project carried out well.

For example, the government could convene the elements of the medical community and sponsor development of guidelines in cooperation with the medical profession. The government could do this through an intervening body such as the Institute of Medicine, the Joint Commission for the Accreditation of Healthcare Organizations, or a consortium of academic health centers. The scientific traditions of the NIH and its perceived distance from the government as payer would make it possible for NIH to play this role. A respected intervening body could bring together the several elements of the medical profession, and might more easily gain their full cooperation than could a government agency like HCFA. The cosponsor or intervening body could also ensure that the work was done well, and provide the

accountability that the federal government would require.[3]

WOULD PRACTICE GUIDELINES IMPROVE MEDICAL PRACTICE?

It is not clear that practice guidelines would improve medical practice. There is little evidence that existing guidelines have been used, or even that most physicians are aware that they exist. However, several factors suggest that guidelines could be more influential in the future:

- It is likely that the better the guidelines, the greater the probability they will be used. The best of the existing guidelines have been available only for a short time, and even they have not been fully evaluated so they have not been demonstrated to be "good" guidelines.

- Physicians have had few incentives to use guidelines. If they are given good guidelines and incentives to use them, they may do so. For example, if sound guidelines were also the basis for utilization and quality review, physicians would have incentive to become familiar with and use the guidelines.

Appendix 4 presents and discusses several arguments that guidelines may not work. Dr. Brook's paper "Practice Guidelines and Practicing Medicine: Are They Compatible" also discusses how guidelines could help physicians and why they are more needed now than in the past.

CONCLUSION

An expanded clinical research base and carefully developed guidelines can play a highly constructive role in the Medicare program. Practice guidelines may be unique among available methods with the potential to slow the increase in expenditures in that they can increase the quality and efficiency of care.

Simply recognizing the potential for practice guidelines will not cause good guidelines to be developed and used. Methods for their development must be perfected, sponsors must be found, and the physician community must lend its expertise and support. Finally, funding must be made available. The federal government has a legitimate role to play in the sponsorship or funding of clinical research and the development of practice guidelines.

---

[3] Independent research and development organizations like RAND or consortia of academic medical centers might function as advisors under contract to the government or through this mechanism to lend expertise, particularly in the early stages of developing and refining methods.

The Conference on Practice Guidelines is designed to discuss how guidelines should be developed and used, and by whom, and what role might be most appropriate for the federal government. We expect the conference to help the Commission and others develop a sound strategy for the development and use of guidelines.

REFERENCES

American College of Cardiology/American Heart Association Task Force on Assessment of Cardiovascular Procedures, "Guidelines for coronary angiography," Journal of the American College of Cardiology 10(4):935-50, 1987

Brook, R.H., M.R. Chassin, et al, "A method for detailed assessment of the appropriateness of medical technologies," International Journal of Technology Assessment in Health Care 2(1):53-63, 1986

Chassin, M.R., J. Kosecoff, et al, "Indications for Selected Medical and Surgical Procedures: A Literature Review and Ratings of Appropriateness: Coronary Angiography," publication R-3204/1-CWF-HF-HCFA-PMT-RWJ (Santa Monica: The Rand Corporation, 1986a)

Chassin, M.R., R.H. Brook, et al, "Variation in the use of medical and surgical services by the Medicare population," New England Journal of Medicine 314:285-290, 1986b

Chassin, M.R., J. Kosecoff, et al, "How coronary angiography is used: Clinical determinations of appropriateness," Journal of the American Medical Association 258(18):2543-2547, 1987a

Chassin, M.R., J. Kosecoff, et al. "Does inappropriate use explain geographic variations in the use of health care services?" Journal of the American Medical Association 258:2533-2537, 1987b

Eddy, D.M,, "Variations in physician practice: The role of uncertainty," Health Affairs 3(2):74-89, 1984

Eisenberg, J.M., Doctors' Decisions and the Cost of Medical Care. (Ann Arbor: Health Administration Press, 1986)

Graboys, T.B., Headley, A., et al, "Results of a second opinion program for coronary artery bypass graft surgery," Journal of the American Medical Association 258:1611-1614, 1987

Greenfield, S., C.E. Lewis, et al, "Peer review by criteria mapping: Criteria for diabetes mellitus," Annals of Internal Medicine 83(6):761-770, 1975.

Greenfield, S., M.A. Nadler, et al, "The clinical investigation and management of chest pain in an emergency department: Quality assessment by criteria mapping," Medical Care 15(11):898-905, 1977.

Greenspan, A.M., Kay, H.R., et al, "Incidence of unwarranted implantation of permanent cardiac pacemakers in a large medical population," New England Journal of Medicine 318:158-163, 1988

Institute of Medicine, National Academy of Sciences, Assessing Medical Technologies (Washington: National Academy Press, 1985)

Mullen, F, and I. Jacoby, "The town meeting for technology: The maturation of Consensus Conferences," Journal of the American Medical Association 254:1068-1072, 1985.

Perry, S, "The NIH Consensus Development Program: A decade later," New England Journal of Medicine 317:485-488, 1987.

Schroeder, S.A., "Strategies for reducing medical costs by changing physicians' behavior: Efficacy and impact on quality of care," International Journal of Technology Assessment in Health Care 3:39-50, 1987

Stulbarg, M.S., B. Gerbert, et al, "Outpatient treatment of chronic obstructive pulmonary disease -- A practitioner's guide," Western Journal of Medicine 142:842-846, 1985.

Wennberg, J.E. and A. Gittelsohn, "Variations in medical care among small areas," Scientific American 246:120-133, 1982

Wennberg, J.E., "Dealing with medical practice variations: A proposal for action," Health Affairs 3:6-32, 1984

Wennberg, J.E., "On patient need, equity, supplier-induced demand and the need to assess the outcome of common medical practices," Medical Care 23:512-520, 1985

Wennberg, J.E., "The paradox of appropriate care," Journal of the American Medical Association 258(18):2568-2569, 1987

Winslow, C.M., Kosecoff, J.B., et al, "The appropriateness of performing coronary artery bypass surgery," Journal of the American Medical Association 260:505-509, 1988a

Winslow, C.M., D.H. Solomon, et al, "The appropriateness of carotid endarterectomy," New England Journal of Medicine 318:721-727, 1988b

APPENDIX I: WHAT ARE GUIDELINES?

Practice guidelines are standardized specifications for care, either for the use of a particular service or procedure or for the management of a specific clinical problem. To be clinically useful, guidelines must be specificied in sufficient detail that they discriminate between what is and what is not appropriate care. To be clinically credible, they must be derived from evidence of effectiveness.

We use the term practice guideline to mean formally developed guidelines based on the clinical research literature and the collective judgments of expert physicians. Such guidelines provide a way to assemble and clarify existing knowledge, to extend it through consensus of expert physicians, and to make it available in a usable form.

The purpose of guidelines is to improve the quality and efficiency of patient care by insuring that diagnostic and therapeutic services provided are both effective and appropriate for the individual patient. Guidelines can do this in two ways: by providing help to individual decisionmakers such as physicians, and as a source of criteria or standards for review by those who would evaluate or influence the process of care.

Guidance for decisionmakers Practice guidelines provide guidance to physicians and others who must make decisions about the use of specific medical procedures in the face of uncertainty. They can guide physicians and patients making decisions for the care of individual patients; health care managers deciding what resources will be available in a hospital or HMO; and administrators making coverage and benefit design decisions.

Informal guidelines have been used by physicians for years. As Eddy points out, physicians follow informal guidelines or "clinical policies" in making many clinical decisions (Eddy, 1983). There is a need for more explicit guidelines providing specific recommendations for care, and developed through a rigorous, structured process.

Criteria or standards for review Practice guidelines can also be used as the clinical basis for utilization review and quality assessment, and to determine eligibility for payment. Guidelines that summarize our best knowledge of the appropriate use of medical services should be the basis for the criteria and standards by which care is judged. If reviewers are to influence physicians' decisions, they should base the review on the best medical knowledge available.

Physicians do not often welcome external review, in part because they do not consider the review to consistently lead to accurate

judgments of the quality of the care. Review criteria and standards based on sound practice guidelines offer a means of improving the accuracy of review.

Types of guidelines with examples

It is useful to distinguish two kinds of practice guidelines. We provide an example of each. The first focuses on a particular procedure or service, such as coronary artery bypass surgery, and provides recommendations for how to use it. The second focuses on a particular patient problem, such as chest pain, and provides recommendations for how to care for a patient with that problem, including which services to provide. These differences are reflected in how the guidelines are used.

Guidelines for a procedure: indications for coronary angiography

The first type specifies when it is appropriate to use the procedure or service. A physician can determine whether the characteristics of a particular patient correspond to appropriate indications for the procedure. Similarly, a utilization review program could use the guidelines either prospectively or retrospectively to identify cases with apparent inappropriate indications for further evaluation. An example follows.

Guidelines developed for the RAND Health Services Utilization Study specify which indications for the use of coronary angiography would be appropriate, inappropriate, or equivocal (Chassin et al, 1987). According to these guidelines, it is appropriate to perform coronary angiography for a patient with nonspecific chest pain and a positive treadmill test, and inappropriate for a patient with stable mild angina who has a negative treadmill test result or has not had a treadmill test:

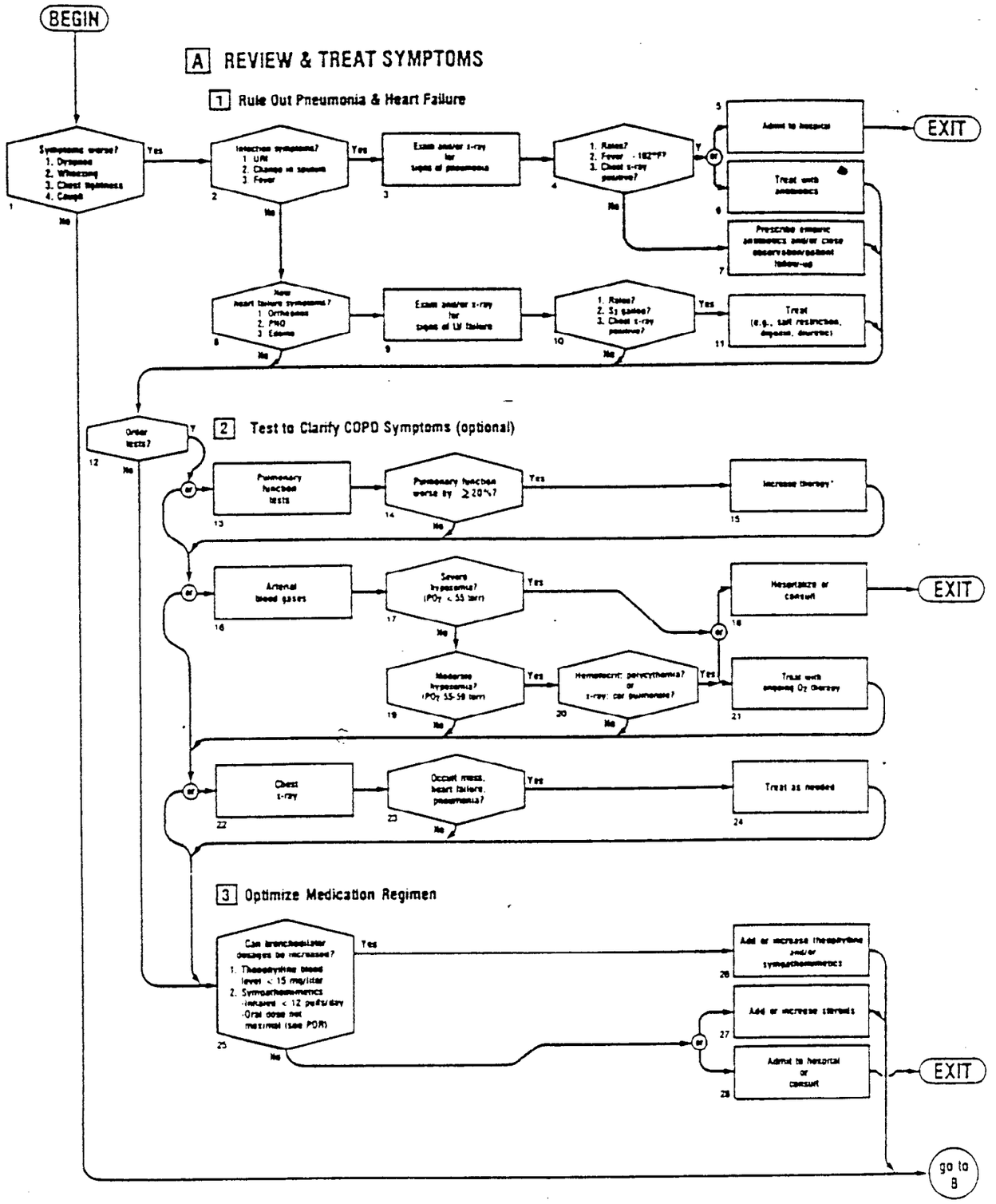
> Coronary angiography is indicated in a patient with chronic stable angina (without strong contraindications to CABG surgery) in whom angina occurs with mild exertion (Class III or IV) and who has received maximal medical management and a very positive exercise ECG and (has) a negative exercise thallium scan and a positive exercise MUGA.
>
> Coronary angiography is not indicated in a women 50 years or younger with chest pain of uncertain origin, no exercise ECG, no exercise thallium scan, and no exercise MUGA.

These guidelines specify in detail the clinical findings needed to guide patient care decisions; they can also be the basis of criteria for evaluation of the appropriateness of use of the procedure. For example, criteria for review could be derived from this set of indications by setting as a standard of acceptable care all indications except those that are considered inappropriate.

Guidelines for care of a patient problem: an algorithms for chronic lung disease The second type of guideline defines which services and what care are appropriate for a particular kind of patient. In their more rigid forms, they are called protocols. These guidelines are more complex because they must describe and recommend action on a number of clinical decisions. Because of the sequential nature of many clinical decisions - the next step may depend on the results of a diagnostic test or response to therapy - these guidelines must include a branching tree of sequential contingencies. They may be more difficult to develop, and to use.

The guidelines developed by Stulbarg and colleagues for outpatient treatment of chronic obstructive pulmonary disease provide a good example (Stulbarg et al, 1985). A portion of this algorithm is attached.

Guidelines like these have been used to develop "criteria maps" to review care of diabetes mellitus, chest pain, and other conditions (Greenfield et al, 1975; Greenfield et al, 1977).



FEV_1 = forced expiratory volume in one second; LV = left ventricular; *PDR* = *Physician's Desk Reference*; PND = paroxysmal nocturnal dyspnea; q.o.d. = every other day; URI = upper respiratory (tract) infection

Figure 1.—A "criteria map" for chronic obstructive pulmonary disease (COPD). The map deals with four major areas of outpatient care in case of COPD: A, Review and treatment of symptoms; B, Review of medications and their side effects; C, Laboratory monitoring; D, Prophylaxis and education. The criteria map has a BEGIN point. EXIT points when a patient enters the hospital or is sent to a chest physician for consultation and a STOP when the four sections of the maps have been traversed. Hexagonal shapes represent possible signs and symptoms of COPD patients; rectangles indicate subsequent physician actions. Arrows direct one from signs or symptoms to physician actions through each section and from one section to the subsequent section. When several alternative actions are acceptable, or signs appear in the map.

APPENDIX 2: GOOD GUIDELINES AND GOOD METHODS

Guidelines and the review criteria derived from them will be more effective in improving care if they are perceived by decisionmakers to be of high quality. Desirable characteristics of guidelines and then of methods are described. Several existing guidelines and the methods used to develop them are then described with reference to these characteristics.

Desirable characteristics of guidelines

- They should be comprehensive, including all likely indications for the use of the procedure (or all decisions to be made in care of a patient problem). They should include all major relevant additional factors that must be taken into consideration in the decision to recommend a procedure.

- They should be specific, clearly describing the exact conditions for which the procedure is recommended.

- They should describe in meaningful detail the distinguishing features that separate one indication from another.

- They should clearly indicate both under what circumstances a procedure is appropriate, and when it is inappropriate.

- They should be inclusive of all major relevant additional factors that must be taken into consideration in the decision to recommend a procedure.

- They should be manageable: the number and complexity of indications must not be so great as to be too difficult to use. Guidelines must be presented in a form and in language that makes them easy to understand and implement.

Desirable characteristics of a method

Practice guidelines should be developed through a rigorous, structured process that synthesizes the information in the scientific literature, extends it through the knowledge of expert physicians, and expresses the information in specific, precise, comprehensive guidelines. In this section, we describe what characteristics such a method should have, and compare several existing methods to the ideal. In the next section we will discuss a related topic: who should sponsor the development of guidelines.

Ideally, guidelines for the use of services should be based entirely on scientific evidence of their effectiveness. The recognized "gold standard" for determining effectiveness is the

randomized clinical trial (RCT). Because such trials are expensive and difficult to carry out, most diagnostic and therapeutic procedures including the majority of the frequently-performed procedures, have not been subjected to RCTs or other comparably rigorous evaluation.

In the absence of conclusive scientific evidence, investigators have turned to less precise methods, such as unstructured review of the literature, or structured synthesis of evidence using meta-analysis or decision analysis. Further, where solid evidence is lacking, clinical experience can sometimes fill in or extend. Practice guidelines represent a marriage of evidence and clinical experience.

A suitable method for synthesizing evidence and extending it through expert clinical opinion to develop practice guidelines must be credible and acceptable to those who would use it, and feasible. A method's credibility and acceptability depend on the following characteristics:

- Does it aggregate or synthesize the best available information on effectiveness from review and synthesis of scientific data from literature?
- Does it define indications that are specific and unambiguous?
- Does the method used to incorporate expert opinion minimize bias? How are the experts chosen? Do the selection and the process used minimize bias and dominance?
- Is the method reliable? Does it produce reproducible results?
- Has it been validated by acceptable methods? Do the results of the process have validity when compared with other evidence?

Feasibility - is the method practical, reproducible, and not too expensive?

- Can the process be carried out by respected professionals?
- Does the method produce meaningful, acceptable guidelines that can be easily understood and used in practice?
- Can the method be implemented now for a large number of procedures at reasonable cost and expenditure of human capital?

Existing guidelines and methods

A number of guidelines have already been developed, most by medical societies and associations. However, the important clinical variables are rarely specified explicitly enough to readily discriminate between care that is appropriate and that which is not.

Most of these have been detailed lists of tests or procedures to be consider obtaining for a particular patient, without guidance as to how or when to use them. Such lists provide little guidance to the decision making process of physicians or reviewers since they do not distinguish between appropriate and inappropriate use and are not considered here.

Several existing guidelines more nearly fulfill the criteria. Similarly, each method used to date to develop guidelines also falls short of fully meeting the ideal. Several methods have many of the characteristics desired, including those used by the American College of Physicians, the American College of Cardiology and American Heart Association, and the method developed by the RAND investigators for the Health Services Utilization Study. Several of these guidelines and methods are described below.

The Clinical Efficacy Assessment Project (CEAP) CEAP began in 1976 as the Medical Necessity Project, a joint undertaking of the American College of Physicians (ACP) and the Blue Cross and Blue Shield (BC/BS) organization. Its purpose was "to identify outmoded tests thereby eliminating reimbursements for useless medical procedures." Since 1981 it has been carried out by the ACP for the purpose of elevating the standards of medical practice. A long list of tests and procedures used in internal medicine (but no surgical operations) have been evaluated.

The major shortcoming of many of the CEAP guidelines is their lack of specificity. While recommendations are comprehensive in scope, many do not clearly specify the clinical differences that determine whether a test or procedure is or is not appropriate for a given individual. For example, the recommendations for lithotripsy do not describe under which conditions, if any, operative removal of stones is preferable to lithotripsy (ref). Duration of symptoms, size of stone, location, comorbidity, initial versus recurrence, anesthetic risk, underlying etiology, renal status, and other factors that might enter into the decision are not discussed. It is difficult, therefore, to apply these guidelines to the individual case, or for a reviewer to determine whether lithotripsy in any given individual was or was not appropriate.

The method used by the American College of Physicians (ACP) to develop these comprehensive guidelines is a fairly rigorous one.

The best available scientific information is gathered by a comprehensive literature review, and recommendations for use are derived by expert consensus with review and approval by interested parties and, finally, by the Board of Regents of the ACP.

However, the group consensus method employs almost entirely internists. For example, the group approving the recommendations on lithotripsy did not include a urologist. The method of review, revision and approval involves further input from specialist and enthusiasts, adding further concern about possible bias. Any method that includes only the judgments of physicians of a single specialty, particularly the specialty that controls the process, is likely to be suspected of generating guidelines that are not free from bias. This is particularly so if physicians from that specialty benefit financially from use of the services in question, and if developers or champions of the service or procedure are included in the physician panels. The chance of bias can be reduced if the panels are chosen carefully to include a wide variety of physicians including both proponents and opponents, if the process of decisionmaking by the panel is carefully structured is open to all observers, and if the results are widely reviewed.

No studies have been presented which demonstrate the reliability or validity of the CEAP method. The CEAP method is clearly feasible. It has been implemented for many services.

Overall, the CEAP method is the most comprehensive effort to date, representing a level of commitment to standards by the ACP that is exemplary. However, with the exception of the probable elimination of use of those tests and procedures which it has found to be worthless (a contribution of no small importance), the effect which will be achieved with CEAP guidelines on the quality of patient care is unclear.

American College of Cardiology and the American Heart Association guidelines. The guidelines generated by the Task Force on Assessment of Diagnostic and Therapeutic Cardiovascular Procedures of American College of Cardiology and the American Heart Association, such as the recently published Guidelines for Percutaneous Transluminal Coronary Angioplasty (ACC/AHA Task Force, 198x), are highly specific and place each indication into one of three classes: Class I, where there is general agreement that the procedure is justified, Class II, where there is divergence of opinion, and Class III, where there is general agreement that the procedure is not indicated. These guidelines specify the details of the clinical and laboratory findings which discriminate among candidates in a meaningful way.

These guidelines are specific and detailed. They clearly indicate when the procedure is appropriate and when it is

inappropriate. The format makes them readily accessible for clinical use. The guidelines are easily understood and can, therefore, be applied readily in practice.

However, the method and completeness of the reviews of scientific information is not explicitly stated. The validity and reproducibility of the results of the expert panel have not been established, and, most importantly, since the panel is, apparently, all cardiologists, questions arise as to its objectivity. The method of decisionmaking by the physician panels is not described.

Acceptability of the ACC-AHA guidelines hinges on the credibility question. The sponsoring organizations are well respected, though open to the challenge of self-interest, and they have indicated their plans for updating and revision. The method is clearly feasible. The ACC-AHA Task Force has applied the method to several procedures.

RAND Health Services Utilization Study As part of the Health Services Utilization Study of the extent and causes of geographic variations in the use of services, the RAND Corporation developed a modified Delphi and interactive group technique for obtaining expert consensus on the appropriateness of indications for six operations and procedures. Highly specific and mutually exclusive indications were derived after comprehensive literature review and consultations with experts. These indications were then rated for appropriateness by panels that were carefully balanced with medical and surgical specialists and generalists (Park, 1986).

The RAND guidelines are the most detailed, comprehensive, and the most specific of any that have been developed. Each of the comprehensive set of indications is specifically and unambiguously defined, with consideration of the relevant clinical variables. The resulting guidelines clearly distinguish between appropriate and inappropriate indications, as well as an intermediate "indeterminate" category.

This method has several excellent features. A comprehensive literature review elicits the relevant scientific information. The method of selecting experts appears to minimize bias. Initial tests of validity are encouraging, but the reliability of the method has not been established.

Questions have been raised about the feasibility of implementing the highly specific RAND methodology. While it is probable that the highly specific indications and ratings can be translated into clinically usable tools, this has not yet been done. Further, no organization has indicated its interest and willingness to commit the resources that would be required to apply this method to the evaluation of a significant number of

operations and procedures.

Consensus Development Project of the National Institutes of Health Sponsored by the Office of Medical Applications of Research, National Institutes of Health (NIH), the Consensus Conferences bring together a prestigious panel of bioscientists and nonmedical experts to hear evidence and testimony in open forum concerning the efficacy of new technologies. The panel then adjourns to write its draft report in private. The draft report is considered during a second hearing for commentary, and the panel then writes a final revised report.

Although discussions are erudite and often fairly thorough, most of the recommendations are general in nature and are not framed in a format that permits discrimination among similar candidates for a procedure. No attempt is made to be all-inclusive in considering indications for the procedure. The report from a recent conference on kidney stones, for example, states that the efficacy and safety of lithotripsy "compares well with that of percutaneous nephrolithotomy" without spelling out in which situations one of these forms of treatment is appropriate and the other is not, or where it is as yet undecided.

NIH consensus statements have been well-regarded, but most are not specific enough to be suitable for implementation as standards.

The major criticisms of the NIH Consensus Conferences relevant to the credibility of the conclusions pertain to the hearing format of the meeting and the requirement that the participants make their recommendations under heavy time pressure.

Diagnostic and Therapeutic Technology Assessment Project of the American Medical Association This AMA project polls a large number (15-120) of specialists for their opinions as to the safety and effectiveness of new technologies. The questions posed are very general, and no attempt is made to be comprehensive in scope. Respondents are asked to classify only safety or effectiveness in a given clinical situation. There are four judgments: established, investigational, indeterminate, or unacceptable.

The usefulness of the answers is limited by the general nature of the questions asked. For example, a judgment on the effectiveness of lithotripsy is requested by the question, "Is transurethral urethroscopy with retrograde manipulation of ureteral stones to the renal pelvis followed by ESWL (Extracorporeal Shock-Wave Lithotripsy) an effective therapy for proximal ureteral stones?" The respondents are given no opportunity to consider other factors that enter into the decision, such as the patient's surgical risk, size of stone, presence of infection, anatomic variations, renal function, etc.

While this question may help determine whether lithotripsy is in general effective and therefore merits coverage, it does not yield the information necessary to determine for which patients lithotripsy is or is not indicated.

The credibility of the DATTA process is limited by the general nature of its questions, and the fact that results represent merely a polling of specialists, rather than a rigorous analysis of the evidence on effectiveness. While this may accurately reflect the current state of opinion, it may not represent the highest level of informed opinion.

APPENDIX 3: SPONSORSHIP

The following criteria for sponsorship are intended to insure that guidelines will be developed in such a way that they are likely to be both useful and used.

- The scope of its activities must be national. Regional societies, local governmental agencies or hospitals are unlikely to be able to muster the necessary resources, nor in most cases be considered sufficiently authoritative to gain acceptability for the results.

- The scope should cover all medical specialties. Alternatively, a coalition or other cooperative arrangement will be needed to insure that all of medical practice is included, and that all interested physicians, regardless of specialty, are included in (or have access to) the process.

- The sponsoring organization must be authoritative and respected for having high ethical and scientific standards. It should be regarded as objective and reasonably free of self-interest, either professional or financial.

- It must be able to draw on the necessary expertise to get the job done. It must have or be able to secure the required resources, including funding.

- It must be willing to develop guidelines for the public domain.

Several organizations have either already engaged in guidelines development or would appear to be logical candidates to do so:

Professional societies and associations Professional societies encompassing all medical specialties would appear to be well-suited to take on the task of developing practice guidelines in several ways. They are national in scope, are respected for having high standards, and can muster the necessary expertise to accomplish the task. Two organizations encompass all, or nearly all, specialties: the AMA, and the Council of Medical Specialty Societies (CMSS). Both have begun activities in guidelines development. Each should have no trouble mustering the professional talent and commitment to accomplish the task.

However, any organization of physicians could be suspected of influencing the guidelines in their own interest. To counter this, a professional organization would have to adopt a method designed to minimize bias, and would have to conduct a process visible to all interested parties. The bias expected of any particular physician group could be balanced by including physicians from a variety of specialties, including those that do

not perform the service in question and so have no financial interest in it.

Specialty societies are similarly qualified, but development of a full range of guidelines would require a coordinating body to insure that guidelines are developed for all of the major services. The primary concern with single specialty society sponsorship is objectivity. While some of the practice guidelines developed by specialty societies have been done well the use of expert panelists recruited entirely from one or even two specialties leaves their conclusions open to questions about objectivity and self-interest. This concern could be addressed by use of better balanced panels in which all relevant specialties and other interested parties are represented.

Specialty societies are may also have difficulty mustering the financial and administrative resources to carry the task of development of practice guidelines to completion and to maintain the guidelines by periodic review and revision as needed. Even for a major organization such as the AMA or the CMSS, the resources required would be formidable.

Federal agencies Federal agencies could undoubtedly draw on the necessary professional expertise to develop guidelines, and would have the resources to accomplish the task. However, the agencies differ considerably in the strength of their tradition of scientific excellence and perceived objectivity.

Because the Health Care Financing Administration (HCFA) is directly concerned with the cost of care, the physician community is unlikely to find them acceptable because of a concern that guidelines would be developed with the objective of cost containment rather than quality of care.

The National Center for Health Services Research (NCHSR) and particularly the National Institutes of Health (NIH) would be perceived as more distant from the financial interests that attach to HCFA. Further, they have a tradition of high standards of scientific objectivity in research, particularly the NIH. For either the NCHSR or the NIH, this work might be complementary to their other work in biomedical and health services research.

The NIH sponsors most of the biomedical research in the U.S. including a number of large scale clinical trials, and has carried out a major program of technology assessment through its Consensus Development Conferences. It could presumably sponsor a larger effort to develop practice guidelines. Several have suggested that a new "National Institute of Health Care Delivery" be established within the NIH for this task. Such an Institute might integrate work to evaluate medical practices as well as develop sound methods to develop practice guidelines, and use the method to do so.

The NCHSR may in the future be responsible for an increased level of clinical research and development. The NCHSR has not mounted an effort to develop practice guidelines, and may not yet have the expertise or the capacity to do so.

Academic health centers Some academic health centers have shown interest in the development of practice guidelines. This activity would appear to be consistent with all three of the classic functions of academic health centers: education, research, and practice at the forefront of medicine. Their standards are respected, they can muster the intellectual resources, and their scientific traditions may minimize questions about objectivity or self-interest. However, even in a consortium, they are unlikely to have the financial resources to complete the task, however, so outside support would be needed. They might be more appropriate as the site to evaluate and refine methods for developing guidelines than for the subsequent application of the method to produce large numbers of guidelines.

The Institute of Medicine (IOM) The prestige of its membership and the well-earned reputation for work of highest scientific caliber make the IOM uniquely well qualified to take responsibility for the development of practice guidelines. In the past, such "production" work has traditionally been outside of its concept of the scope of its activities, however. The IOM may also be more appropriate for developing sound methods than for applying them in a large scale effort to produce guidelines.

The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) The avowed purpose of the Joint Commission on the Accreditation of Healthcare Organizations is to set and enforce standards for care. It is clearly national in scope, and is highly respected by both the medical profession and hospital administrators. Few would question its objectivity.

APPENDIX 4: ARGUMENTS AGAINST A STRATEGY BASED ON GUIDELINES

Most doctors do not perceive a need for detailed guidelines or standards. They are not dissatisfied with the quality of care they provide, and from the practitioner's perspective, individualized decision making is more satisfying than use of standardized criteria or algorithms. Many are skeptical that existing methodology is adequate to accomplish the objective of defining quality of care in a useful way. They assert that a strategy to increase quality and efficiency of care built on developing sound practice guidelines will not work.

We will discuss three arguments against the use of practice guidelines:

- The concept is flawed in that it is not possible to develop meaningful guidelines, and those based on opinion in the absence of evidence of effectiveness will solidify untested opinion and do more harm than good.
- Their use degrades the practice of medicine.
- They will be misused.

We will then examine the experience with the use of guidelines to determine the extent to which these criticisms are valid and to gain insight into how the use of guidelines can be improved.

It is not possible to develop meaningful guidelines

The notion that it is not possible to develop meaningful guidelines stems from a respect for the complexity of modern medical practice. Critics argue that practice guidelines cannot deal with all of the important variables in the decision making process. Specifically, practice guidelines cannot realistically take into account the presence of all possible disease states (comorbidity), patient values and preferences, or variability of patient response to treatment. It is also difficult for guidelines to deal with resource availability, the quality of professional performance, and cost effectiveness.

The validity of these concerns relates in large measure to how practice guidelines are used. For indications that are deemed appropriate, few would question the need for the physician to consider all the other relevant factors before making a specific decision or recommendation. On the other hand, if a procedure is ineffective, then individual patient and resource constraints matter little. A responsible physician will not recommend or perform a useless procedure, even if the patient requests it.

A more serious criticism of practice guidelines concerns the

methods by which they are derived. In the absence of good clinical data, such as from a randomized clinical trial, most guidelines are the result of a consensus of expert opinion. They may overrate the value of these opinions. They will reflect the background, biases, and special interests of the panelists, as well as possible dominance by one or two individuals. The practitioner may consider his opinion to be as good as that of experts, particularly if one of the panelists is known to be an outspoken advocate of a particular procedure.

This potential problem must be minimized by using a good method to develop guidelines. In particular, the method must give full weight to the available evidence and must limit the influence of physician opinion to its proper place (see Mr. Eddy's paper "Designing Practice Guidelines" prepared for the conference).

A third criticism relating to feasibility is that in the development of practice guidelines one is dealing with a moving target. Indications for a procedure change as new evidence becomes available. Guidelines can be out of date by the time they are issued. However, indications for most procedures change very little over time. But this argument points to the need for a process for revising guidelines periodically, and for an appeal mechanism as part of implementation.

Their use degrades the practice of medicine

Even if meaningful practice guidelines can be developed, some think their use degrades the practice of medicine by forcing the physician to practice according to a "cookbook". It is argued that requiring the use of guidelines deprofessionalizes medicine by substituting the guidelines for the physician's judgment.

This argument also misperceives the proper role of guidelines in practice. The function of guidelines is to inform the physician and others as to what does and what does not work. Few would argue that worthless procedures should be used. On the other hand, if an operation or procedure is judged to be effective, the physician still must evaluate all the other factors and then help the patient decide what is best. Guidelines are an adjunct to decision making, a means of formalizing, standardizing, and improving the decision making process, not a substitute for it.

Of course, only good guidelines should be used -- a poor guideline may "degrade" medical practice if used at all.

They will be misused by physicians or reviewers

For guidelines to increase the efficiency and quality of care, they must be used properly. Misuse of even the best guidelines can lead to poor care. For example, those applying them must recognize that guidelines cannot anticipate every possible

variation in patients' clinical characteristics and recognize the importance of the attending physician's clinical judgment. Inflexible application of even the best guidelines could result in "cookbook medicine"--and poor clinical decisions.

Many physicians view practice guidelines as but one more instrument in the increasing regulation of the practice of medicine. Few things arouse physician ire as much as the "hassle" surrounding the use of pre-admission screening criteria. While these are usually quite superficial guidelines, their occasional misuse, especially when implemented by non-professionals, increases the concern that payers will tend to implement any system of guidelines blindly. Without a system for dealing with exceptions, it is argued, blanket application of guidelines will be detrimental to patient care.

Payers, on the other hand, are concerned that physicians will "game" the system by describing the indications in such a way that the patient appears to meet the criteria when she doesn't. While this possibility cannot be dismissed, it is unlikely that very many physicians would respond in this manner, or that they would be able to do it very long if there is effective peer review of performance.

Guidelines haven't worked.

Finally, the most telling argument against the use of practice guidelines is the lack of evidence that the use of guidelines improves the quality of patient care. While the value of guidelines would appear to some to be intuitively obvious - for example, eliminating inappropriate procedures must improve quality of care - there are few, if any, prospective controlled studies that demonstrate that the health status of a population is improved by implementation of guidelines or standards of care in the health care system. In part this is because of the poor quality of most guidelines that have been used, in part it is because of the failure to use them at all.

Guidelines are only effective if they are integrated into the decision making process by the physician in his daily practice. There are few instances where that has been occurred. This seems to be true even when sponsored by an organization that would be expected to have significant influence because of its prestige or membership.

For example, the American College of Obstetricians and Gynecologists (ACOG) issued guidelines for performance of cesarean section delivery that unequivocally declared that prior cesarean-section was not an appropriate indication for another. Several years later, a survey showed that a substantial fraction of obstetricians still performed cesarean-section for that reason, and still considered it to be a valid indication ( ).

The Community Cancer Program Breast Study experience was similar. Oncology leaders in each hospital participated in the derivation of standards for management of patients with breast cancer. A key feature was the necessity of recording the stage of disease for each patient prior to therapy. Followup evaluation not only revealed than many physicians failed to use staging, but that the hospital study leaders, and even the developers of the program failed to meet the standard in a significant fraction of patients ( ).

Studies also suggest that the impact of the NIH Consensus conferences has been small. Even in regions where there was familiarity with the recommendation of the conference, practice patterns did not change (Kanouse et al, 1987).

Some of the problem may be related to the quality of the practice guidelines that are used. If guidelines are simplistic, physicians will ignore them; if they are too general or nonspecific, they provide no real guidance.

Another reason guidelines haven't been readily adopted is that some physicians consider them to be an infringement on their sovereignty as arbiters of the patient's welfare. While doctors necessarily use self-generated guidelines in daily practice, they do not readily accept imposition of standards from outside. In a sense this resistance results from confusion about the intent of guidelines. With or without guidelines, it remains the physician's unquestioned responsibility is to consider all relevant factors and then recommend a specific treatment for the individual patient. But it is clearly not feasible for any individual physician to carry out a rigorous evaluation of the effectiveness of all treatments under all circumstances. Good guidelines in effect summarize that evidence -- having that done by experts using a structured process should help the physician, and does not diminish the role or responsibility of the individual physician to choose the best possible treatment for his patient.

Summary

These arguments are not without merit. They point out the importance of proper development and proper use of guidelines. Sound practice guidelines can be developed. Valid criteria or standards for review can be derived from them. If properly derived, sponsored, and implemented, they have the potential to improve care by eliminating the use of ineffective and outdated procedures. Nevertheless, even the best criteria can be misapplied, so that systems for development of guidelines must be coupled with appropriate mechanisms for their use.

Practice guidelines have a positive character that is consonant

with the goals of physicians to provide efficient care of good quality. Unlike utilization review and similar approaches to reducing unnecessary services, guidelines also specify the right way to provide care. Thus they provide physicians an opportunity to take greater responsibility for the quality and efficiency of care.