

# The Quality of Care

## How Can It Be Assessed?

Avedis Donabedian, MD, MPH

Before assessment can begin we must decide how quality is to be defined and that depends on whether one assesses only the performance of practitioners or also the contributions of patients and of the health care system; on how broadly health and responsibility for health are defined; on whether the maximally effective or optimally effective care is sought; and on whether individual or social preferences define the optimum. We also need detailed information about the causal linkages among the structural attributes of the settings in which care occurs, the processes of care, and the outcomes of care. Specifying the components or outcomes of care to be sampled, formulating the appropriate criteria and standards, and obtaining the necessary information are the steps that follow. Though we know much about assessing quality, much remains to be known.

(JAMA 1988;260:1743-1748)

THERE was a time, not too long ago, when this question could not have been asked. The quality of care was considered to be something of a mystery: real, capable of being perceived and appreciated, but not subject to measurement.

For editorial comment see p 1759.

The very attempt to define and measure quality seemed, then, to denature and belittle it. Now, we may have moved too far in the opposite direction. Those who have not experienced the intricacies of clinical practice demand measures that are easy, precise, and complete—as if a sack of potatoes was being weighed.

From the University of Michigan School of Public Health, Ann Arbor.

This article was written for the *AMA Lectures in Medical Science*; it is the basis for a lecture in that series given on Jan 11, 1988, by invitation of the Division of Basic Sciences, American Medical Association, Chicago.

Reprint requests to 1739 Ivywood Dr, Ann Arbor, MI 48103.

True, some elements in the quality of care are easy to define and measure, but there are also profundities that still elude us. We must not allow anyone to belittle or ignore them; they are the secret and glory of our art. Therefore, we should avoid claiming for our capacity to assess quality either too little or too much. I shall try to steer this middle course.

### SPECIFYING WHAT QUALITY IS

#### Level and Scope of Concern

Before we attempt to assess the quality of care, either in general terms or in any particular site or situation, it is necessary to come to an agreement on what the elements that constitute it are. To proceed to measurement without a firm foundation of prior agreement on what quality consists in is to court disaster.<sup>1</sup>

As we seek to define quality, we soon become aware of the fact that several formulations are both possible and legitimate, depending on where we are

located in the system of care and on what the nature and extent of our responsibilities are. These several formulations can be envisaged as a progression, for example, as steps in a ladder or as successive circles surrounding the bull's-eye of a target. Our power, our responsibility, and our vulnerability all flow from the fact that we are the foundation for that ladder, the focal point for that family of concentric circles. We must begin, therefore, with the performance of physicians and other health care practitioners.

As shown in Fig 1, there are two elements in the performance of practitioners: one technical and the other interpersonal. Technical performance depends on the knowledge and judgment used in arriving at the appropriate strategies of care and on skill in implementing those strategies. The goodness of technical performance is judged in comparison with the best in practice. The best in practice, in its turn, has earned that distinction because, on the average, it is known or believed to produce the greatest improvement in health. This means that the goodness of technical care is proportional to its expected ability to achieve those improvements in health status that the current science and technology of health care have made possible. If the realized fraction of what is achievable is called *effectiveness*, the quality of technical care becomes proportionate to its effectiveness (Fig 2).

Here, two points deserve emphasis. First, judgments on technical quality are contingent on the best in current knowledge and technology; they cannot go beyond that limit. Second, the judg-

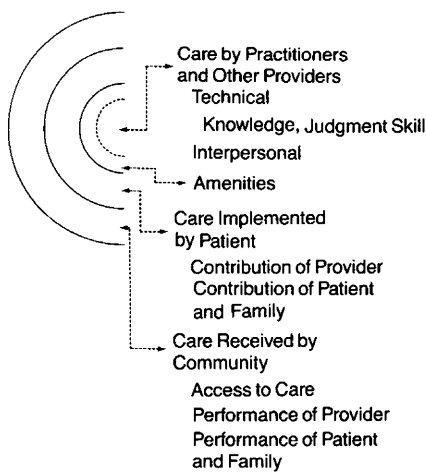


Fig 1.—Levels at which quality may be assessed.

ment is based on future expectations, not on events already transpired. Even if the actual consequences of care in any given instance prove to be disastrous, quality must be judged as good if care, at the time it was given, conformed to the practice that could have been expected to achieve the best results.

The management of the interpersonal relationship is the second component in the practitioner's performance. It is a vitally important element. Through the interpersonal exchange, the patient communicates information necessary for arriving at a diagnosis, as well as preferences necessary for selecting the most appropriate methods of care. Through this exchange, the physician provides information about the nature of the illness and its management and motivates the patient to active collaboration in care. Clearly, the interpersonal process is the vehicle by which technical care is implemented and on which its success depends. Therefore, the management of the interpersonal process is to a large degree tailored to the achievement of success in technical care.

But the conduct of the interpersonal process must also meet individual and social expectations and standards, whether these aid or hamper technical performance. Privacy, confidentiality, informed choice, concern, empathy, honesty, tact, sensitivity—all these and more are virtues that the interpersonal relationship is expected to have.

If the management of the interpersonal process is so important, why is it so often ignored in assessments of the

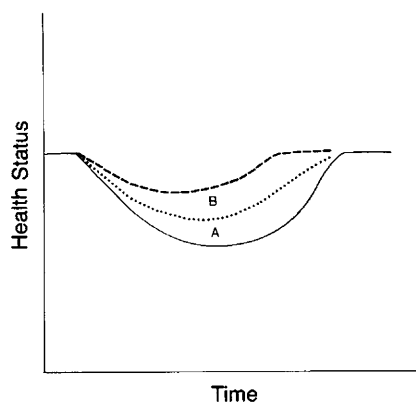


Fig 2.—Graphical presentation of effectiveness (in a self-limiting disease). Solid line indicates course of illness without care; dotted line, course of illness with care to be assessed; and dashed line, course of illness with "best" care. Effectiveness equals  $A/(A+B)$ .

quality of care? There are many reasons. Information about the interpersonal process is not easily available. For example, in the medical record, special effort is needed to obtain it. Second, the criteria and standards that permit precise measurement of the attributes of the interpersonal process are not well developed or have not been sufficiently called upon to undertake the task. Partly, it may be because the management of the interpersonal process must adapt to so many variations in the preferences and expectations of individual patients that general guidelines do not serve us sufficiently well.

Much of what we call the *art of medicine* consists in almost intuitive adaptations to individual requirements in technical care as well as in the management of the interpersonal process. Another element in the art of medicine is the way, still poorly understood, in which practitioners process information to arrive at a correct diagnosis and an appropriate strategy of care.<sup>2</sup> As our understanding of each of these areas of performance improves, we can expect the realm of our science to expand and that of our art to shrink. Yet I hope that some of the mystery in practice will always remain, since it affirms and celebrates the uniqueness of each individual.

The science and art of health care, as they apply to both technical care and the management of the interpersonal process, are at the heart of the metaphorical family of concentric circles depicted in Fig 1. Immediately surrounding the center we can place the

amenities of care, these being the desirable attributes of the settings within which care is provided. They include convenience, comfort, quiet, privacy, and so on. In private practice, these are the responsibility of the practitioner to provide. In institutional practice, the responsibility for providing them devolves on the owners and managers of the institution.

By moving to the next circle away from the center of our metaphorical target, we include in assessments of quality the contributions to care of the patients themselves as well as of members of their families. By doing so we cross an important boundary. So far, our concern was primarily with the performance of the providers of care. Now, we are concerned with judging the care as it actually was. The responsibility, now, is shared by provider and consumer. As already described, the management of the interpersonal process by the practitioner influences the implementation of care by and for the patient. Yet, the patient and family must, themselves, also carry some of the responsibility for the success or failure of care. Accordingly, the practitioner may be judged blameless in some situations in which the care, as implemented by the patient, is found to be inferior.

We have one more circle to visit, another watershed to cross. Now, we are concerned with care received by the community as a whole. We must now judge the social distribution of levels of quality in the community.<sup>3</sup> This depends, in turn, on who has greater or lesser access to care and who, after gaining access, receives greater or lesser qualities of care. Obviously, the performance of individual practitioners and health care institutions has much to do with this. But, the quality of care in a community is also influenced by many factors over which the providers have no control, although these are factors they should try to understand and be concerned about.

I have tried, so far, to show that the definition of quality acquires added elements as we move outward from the performance of the practitioners, to the care received by patients, and to the care received by communities. The definition of quality also becomes narrower or more expansive, depending on how narrowly or broadly we define the concept of health and our responsibility for it. It makes a difference in the assessment of our performance whether we see ourselves as responsible for bringing about improvements only in specific aspects of physical or physiological function or whether we include psychological and social function as well.

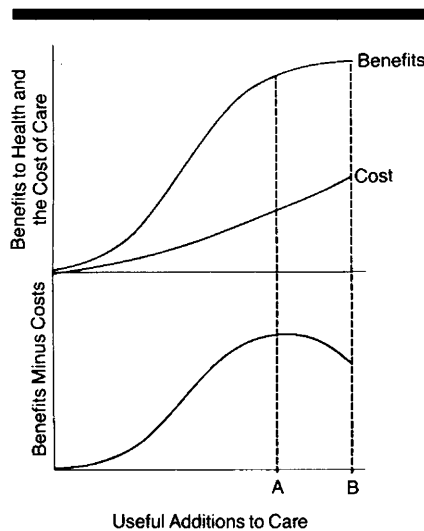


Fig 3.—Hypothetical relations between health benefits and cost of care as useful additions are made to care. A indicates optimally effective care; and B, maximally effective care.

### Valuation of the Consequences of Care

Still another modification in the assessment of performance depends on who is to value the improvements in health that care is expected to produce. If it is our purpose to serve the best interest of our patients, we need to inform them of the alternatives available to them, so they can make the choice most appropriate to their preferences and circumstances. The introduction of patient preferences, though necessary to the assessment of quality, is another source of difficulty in implementing assessment. It means that no preconceived notion of what the objectives and accomplishments of care should be will precisely fit any given patient. All we can hope for is a reasonable approximation, one that must then be subject to individual adjustment.<sup>4,6</sup>

### Monetary Cost as a Consideration

Finally, we come to the perplexing question of whether the monetary cost of care should enter the definition of quality and its assessment.<sup>17</sup> In theory, it is possible to separate quality from inefficiency. Technical quality is judged by the degree to which achievable improvements in health can be expected to be attained. Inefficiency is judged by the degree to which expected improvements in health are achieved in an unnecessarily costly manner. In practice, lower quality and inefficiency coexist because wasteful care is either directly harmful to health or is harmful by displacing more useful care.

Cost and quality are also confounded because, as shown in Fig 3, it is believed that as one adds to care, the corresponding improvements in health become progressively smaller while costs continue to rise unabated. If this is true, there will be a point beyond which additions to care will bring about improvements that are too small to be worth the added cost. Now, we have a choice. We can ignore cost and say that the highest quality is represented by care that can be expected to achieve the greatest improvement in health; this is a "maximalist" specification of quality. Alternatively, if we believe that cost is important, we would say that care must stop short of including elements that are disproportionately costly compared with the improvements in health that they produce. This is an "optimalist" specification of quality. A graphical representation of these alternatives is shown in Fig 3.

Health care practitioners tend to prefer a maximalist standard because they only have to decide whether each added element of care is likely to be useful. By contrast, the practice of optimal care requires added knowledge of costs, and also some method of weighing each added bit of expected usefulness against its corresponding cost.<sup>8</sup> Yet, the practice of optimal care is traditional, legitimate, even necessary, as long as costs and benefits are weighed jointly by the practitioner and the fully informed patient. A difficult, perhaps insoluble, problem arises when a third party (for example, a private insurer or a governmental agency) specifies what the optimum that defines quality is.<sup>9</sup>

### Preliminaries to Quality Assessment

Before we set out to assess quality, we will have to choose whether we will adopt a maximal or optimal specification of quality and, if the latter, whether we shall accept what is the optimum for each patient or what has been defined as socially optimal. Similarly, we should have decided (1) how health and our responsibility for it is to be defined, (2) whether the assessment is to be of the performance of practitioners only or also include that of patients and the health care system, and (3) whether the amenities and the management of the interpersonal process are to be included in addition to technical care. In a more practical vein, we need to answer certain questions: Who is being assessed? What are the activities being assessed? How are these activities supposed to be conducted? What are they meant to accomplish? When we agree on the answers to these questions we are ready to look for the measures that will give

us the necessary information about quality.

### Approaches to Assessment

The information from which inferences can be drawn about the quality of care can be classified under three categories: "structure," "process," and "outcome."<sup>9,10</sup>

**Structure.**—Structure denotes the attributes of the settings in which care occurs. This includes the attributes of material resources (such as facilities, equipment, and money), of human resources (such as the number and qualifications of personnel), and of organizational structure (such as medical staff organization, methods of peer review, and methods of reimbursement).

**Process.**—Process denotes what is actually done in giving and receiving care. It includes the patient's activities in seeking care and carrying it out as well as the practitioner's activities in making a diagnosis and recommending or implementing treatment.

**Outcome.**—Outcome denotes the effects of care on the health status of patients and populations. Improvements in the patient's knowledge and salutary changes in the patient's behavior are included under a broad definition of health status, and so is the degree of the patient's satisfaction with care.

This three-part approach to quality assessment is possible only because good structure increases the likelihood of good process, and good process increases the likelihood of a good outcome. It is necessary, therefore, to have established such a relationship before any particular component of structure, process, or outcome can be used to assess quality. The activity of quality assessment is not itself designed to establish the presence of these relationships. There must be preexisting knowledge of the linkage between structure and process, and between process and outcome, before quality assessment can be undertaken.

Knowledge about the relationship between structure and process (or between structure and outcome) proceeds from the organizational sciences. These sciences are still relatively young, so our knowledge of the effects of structure is rather scanty.<sup>11,12</sup> Furthermore, what we do know suggests that the relationship between structural characteristics and the process of care is rather weak. From these characteristics, we can only infer that conditions are either inimical or conducive to good care. We cannot assert that care, in fact, has been good or bad. Structural characteristics should be a major preoccupation in system design; they are a rather blunt in-

strument in quality assessment.

As I have already mentioned, knowledge about the relationship between attributes of the interpersonal process and the outcome of care should derive from the behavioral sciences. But so far, these sciences have contributed relatively little to quality assessment. I cannot say whether this is because of a deficiency in these sciences or a narrowness in those who assess quality.

Knowledge about the relationship between technical care and outcome derives, of course, from the health care sciences. Some of that knowledge, as we know, is pretty detailed and firm, deriving from well-conducted trials or extensive, controlled observations. Some of it is of dubious validity and open to question. Our assessments of the quality of the technical process of care vary accordingly in their certainty and persuasiveness. If we are confident that a certain strategy of care produces the best outcomes in a given category of patients, we can be equally confident that its practice represents the highest quality of care, barring concern for cost. If we are uncertain of the relationship, then our assessment of quality is correspondingly uncertain. It cannot be emphasized too strongly that our ability to assess the quality of technical care is bounded by the strengths and weaknesses of our clinical science.

There are those who believe that direct assessment of the outcome of care can free us from the limitations imposed by the imperfections of the clinical sciences. I do not believe so. Because a multitude of factors influence outcome, it is not possible to know for certain, even after extensive adjustments for differences in case mix are made, the extent to which an observed outcome is attributable to an antecedent process of care. Confirmation is needed by a direct assessment of the process itself, which brings us to the position we started from.

The assessment of outcomes, under rigorously controlled circumstances, is, of course, the method by which the goodness of alternative strategies of care is established. But, quality assessment is neither clinical research nor technology assessment. It is almost never carried out under the rigorous controls that research requires. It is, primarily, an administrative device used to monitor performance to determine whether it continues to remain within acceptable bounds. Quality assessment can, however, make a contribution to research if, in the course of assessment, associations are noted between process and outcome that seem inexplicable by current knowledge.

Such discrepancies would call for elucidation through research.

If I am correct in my analysis, we cannot claim either for the measurement of process or the measurement of outcomes an inherently superior validity compared with the other, since the validity of either flows to an equal degree from the validity of the science that postulates a linkage between the two. But, process and outcome do have, on the whole, some different properties that make them more or less suitable objects of measurement for given purposes. Information about technical care is readily available in the medical record, and it is available in a timely manner, so that prompt action to correct deficiencies can be taken. By contrast, many outcomes, by their nature, are delayed, and if they occur after care is completed, information about them is not easy to obtain. Outcomes do have, however, the advantage of reflecting all contributions to care, including those of the patient. But this advantage is also a handicap, since it is not possible to say precisely what went wrong unless the antecedent process is scrutinized.

This brief exposition of strengths and weaknesses should lead to the conclusion that in selecting an approach to assessment one needs to be guided by the precise characteristics of the elements chosen. Beyond causal validity, which is the essential requirement, one is guided by attributes such as relevance to the objectives of care, sensitivity, specificity, timeliness, and costliness.<sup>10(p106-118)</sup> As a general rule, it is best to include in any system of assessment, elements of structure, process, and outcome. This allows supplementation of weakness in one approach by strength in another; it helps one interpret the findings; and if the findings do not seem to make sense, it leads to a reassessment of study design and a questioning of the accuracy of the data themselves.

Before we leave the subject of approaches to assessment, it may be useful to say a few words about patient satisfaction as a measure of the quality of care. Patient satisfaction may be considered to be one of the desired outcomes of care, even an element in health status itself. An expression of satisfaction or dissatisfaction is also the patient's judgment on the quality of care in all its aspects, but particularly as concerns the interpersonal process. By questioning patients, one can obtain information about overall satisfaction and also about satisfaction with specific attributes of the interpersonal relationship, specific components of technical care, and the outcomes of care. In doing so, it should be remembered that, un-

less special precautions are taken, patients may be reluctant to reveal their opinions for fear of alienating their medical attendants. Therefore, to add to the evidence at hand, information can also be sought about behaviors that indirectly suggest dissatisfaction. These include, in addition to complaints registered, premature termination of care, other forms of noncompliance, termination of membership in a health plan, and seeking care outside the plan.

It is futile to argue about the validity of patient satisfaction as a measure of quality. Whatever its strengths and limitations as an indicator of quality, information about patient satisfaction should be as indispensable to assessments of quality as to the design and management of health care systems.

## SAMPLING

If one wishes to obtain a true view of care as it is actually provided, it is necessary to draw a proportionally representative sample of cases, using either simple or stratified random sampling. Because cases are primarily classified by diagnosis, this is the most frequently used attribute for stratification. But, one could use other attributes as well: site of care, specialty, demographic and socioeconomic characteristics of patients, and so on.

There is some argument as to whether patients are to be classified by discharge diagnosis, admission diagnosis, or presenting complaint. Classification by presenting complaint (for example, headache or abdominal pain) offers an opportunity to assess both success and failure in diagnosis. If discharge diagnoses are used, one can tell if the diagnosis is justified by the evidence; the failure to diagnose is revealed only if one has an opportunity to find cases misclassified under other diagnostic headings.

A step below strictly proportionate sampling, one finds methods designed to provide an illustrative rather than a representative view of quality. For example, patients may be first classified according to some scheme that represents important subdivisions of the realm of health care in general, or important components in the activities and responsibilities of a clinical department or program in particular. Then, one purposively selects, within each class, one or more categories of patients, identified by diagnosis or otherwise, whose management can be assumed to typify clinical performance for that class.

This is the "tracer method" proposed by Kessner and coworkers.<sup>13,14</sup> The validity of the assumption that the cases selected for assessment represent all

cases in their class has not been established.

Most often, those who assess quality are not interested in obtaining a representative, or even an illustrative picture of care as a whole. Their purposes are more managerial, namely, to identify and correct the most serious failures in care and, by doing so, to create an environment of watchful concern that motivates everyone to perform better. Consequently, diagnostic categories are selected according to importance, perhaps using Williamson's<sup>15</sup> principle of "maximum achievable benefit," meaning that the diagnosis is frequent, deficiencies in care are common and serious, and the deficiencies are correctable.

Still another approach to sampling for managerial or reformist purposes is to begin with cases that have suffered an adverse outcome and study the process of care that has led to it. If the outcome is infrequent and disastrous (a maternal or perinatal death, for example), every case might be reviewed. Otherwise, a sample of adverse outcomes, with or without prior stratification, could be studied.<sup>16-18</sup> There is some evidence that, under certain circumstances, this approach will identify a very high proportion of serious deficiencies in the process of care, but not of deficiencies that are less serious.<sup>19</sup>

## MEASUREMENT

The progression of steps in quality assessment that I have described so far brings us, at last, to the critical issue of measurement. To measure quality, our concepts of what quality consists in must be translated to more concrete representations that are capable of some degree of quantification—at least on an ordinal scale, but one hopes better. These representations are the criteria and standards of structure, process, and outcome.<sup>20,21</sup>

Ideally, the criteria and standards should derive, as I have already implied, from a sound, scientifically validated fund of knowledge. Failing that, they should represent the best informed, most authoritative opinion available on any particular subject. Criteria and standards can also be inferred from the practice of eminent practitioners in a community. Accordingly, the criteria and standards vary in validity, authoritativeness, and rigor.

The criteria and standards of assessment can also be either implicit or explicit. Implicit, unspoken criteria are used when an expert practitioner is given information about a case and asked to use personal knowledge and experience to judge the goodness of the process of

care or of its outcome. By contrast, explicit criteria and standards for each category of cases are developed and specified in advance, often in considerable detail, usually by a panel of experts, before the assessment of individual cases begins. These are the two extremes in specification; there are intermediate variants and combinations as well.

The advantage in using implicit criteria is that they allow assessment of representative samples of cases and are adaptable to the precise characteristics of each case, making possible the highly individualized assessments that the conceptual formulation of quality envisaged. The method is, however, extremely costly and rather imprecise, the imprecision arising from inattentiveness or limitations in knowledge on the part of the reviewer and the lack of precise guidelines for quantification.

By comparison, explicit criteria are costly to develop, but they can be used subsequently to produce precise assessments at low cost, although only cases for which explicit criteria are available can be used in assessment. Moreover, explicit criteria are usually developed for categories of cases and, therefore, cannot be adapted readily to the variability among cases within a category. Still another problem is the difficulty in developing a scoring system that represents the degree to which the deficiencies in care revealed by the criteria influence the outcome of care.

Taking into account the strengths and limitations of implicit and explicit criteria, it may be best to use both in sequence or in combination. One frequently used procedure is to begin with rather abridged explicit criteria to separate cases into those likely to have received good care and those not. All the latter, as well as a sample of the former, are then assessed in greater detail using implicit criteria, perhaps supplemented by more detailed explicit criteria.

At the same time, explicit criteria themselves are being improved. As their use expands, more diagnostic categories have been included. Algorithmic criteria have been developed that are much more adaptable to the clinical characteristics of individual patients than are the more usual criteria lists.<sup>22,23</sup> Methods for weighting the criteria have also been proposed, although we still do not have a method of weighting that is demonstrably related to degree of impact on health status.<sup>24</sup>

When outcomes are used to assess the quality of antecedent care, there is the corresponding problem of specifying the several states of dysfunction and of weighting them in importance relative

to each other using some system of preferences. It is possible, of course, to identify specific outcomes, for example, reductions in fatality or blood pressure, and to measure the likelihood of attaining them. It is also possible to construct hierarchical scales of physical function so that any position on the scale tells us what functions can be performed and what functions are lost.<sup>25</sup> The greatest difficulty arises when one attempts to represent as a single quantity various aspects of functional capacity over a life span. Though several methods of valuation and aggregation are available, there is still much controversy about the validity of the values and, in fact, about their ethical implications.<sup>26,27</sup> Nevertheless, such measures, sometimes called *measures of quality-adjusted life*, are being used to assess technological innovations in health care and, as a consequence, play a role in defining what good technical care is.<sup>28,29</sup>

## INFORMATION

All the activities of assessment that I have described depend, of course, on the availability of suitable, accurate information.

The key source of information about the process of care and its immediate outcome is, no doubt, the medical record. But we know that the medical record is often incomplete in what it documents, frequently omitting significant elements of technical care and including next to nothing about the interpersonal process. Furthermore, some of the information recorded is inaccurate because of errors in diagnostic testing, in clinical observation, in clinical assessment, in recording, and in coding. Another handicap is that any given set of records usually covers only a limited segment of care, that in the hospital, for example, providing no information about what comes before or after. Appropriate and accurate recording, supplemented by an ability to collate records from various sites, is a fundamental necessity to accurate, complete quality assessment.

The current weakness of the record can be rectified to some extent by independent verification of the accuracy of some of the data it contains, for example, by reexamination of pathological specimens, x-ray films, and electrocardiographic tracings and by recoding diagnostic categorization. The information in the record can also be supplemented by interviews with, or questionnaires to, practitioners and patients, information from patients being indispensable if compliance, satisfaction, and some long-term outcomes are to be assessed. Sometimes, if more precise

information on outcomes is needed, patients may have to be called back for reexamination. And for some purposes, especially when medical records are very deficient, videotaping or direct observation by a colleague have been used, even though being observed might itself elicit an improvement in practice.<sup>30,31</sup>

## CONCLUSIONS

In the preceding account, I have detailed, although rather sketchily, the steps to be taken in endeavoring to assess the quality of medical care. I hope it is clear that there is a way, a path

worn rather smooth by many who have gone before us. I trust it is equally clear that we have, as yet, much more to learn. We need to know a great deal more about the course of illness with and without alternative methods of care. To compare the consequences of these methods, we need to have more precise measures of the quantity and quality of life. We need to understand more profoundly the nature of the interpersonal exchange between patient and practitioner, to learn how to identify and quantify its attributes, and to determine in what ways these contribute to the patient's health and welfare. Our

information about the process and outcome of care needs to be more complete and more accurate. Our criteria and standards need to be more flexibly adaptable to the finer clinical peculiarities of each case. In particular, we need to learn how to accurately elicit the preferences of patients to arrive at truly individualized assessments of quality. All this has to go on against the background of the most profound analysis of the responsibilities of the health care professions to the individual and to society.

## References

1. Donabedian A: *The Definition of Quality and Approaches to Its Management*, vol 1: *Explorations in Quality Assessment and Monitoring*. Ann Arbor, Mich, Health Administration Press, 1980.
2. Eraker S, Politser P: How decisions are reached: Physician and patient. *Ann Intern Med* 1982;97:262-268.
3. Donabedian A: Models for organizing the delivery of health services and criteria for evaluating them. *Milbank Q* 1972;50:103-154.
4. McNeil BJ, Weichselbaum R, Pauker SG: Fallacy of the five-year survival in lung cancer. *N Engl J Med* 1978;299:1397-1401.
5. McNeil BJ, Weichselbaum R, Pauker SG: Tradeoffs between quality and quantity of life in laryngeal cancer. *N Engl J Med* 1981;305:982-987.
6. McNeil BJ, Pauker SG, Sox HC Jr, et al: On the elicitation of preferences for alternative therapies. *N Engl J Med* 1982;306:1259-1262.
7. Donabedian A, Wheeler JRC, Wyszewianski L: Quality, cost, and health; An integrative model. *Med Care* 1982;20:975-992.
8. Torrance GW: Measurement of health status utilities for economic appraisal: A review. *J Health Econ* 1986;5:1-30.
9. Donabedian A: Quality, cost, and clinical decisions. *Ann Am Acad Polit Soc Sci* 1983;468:196-204.
10. Donabedian A: Evaluating the quality of medical care. *Milbank Q* 1966;44:166-203.
11. Palmer RH, Reilly MC: Individual and institutional variables which may serve as indicators of quality of medical care. *Med Care* 1979;17:693-717.
12. Donabedian A: The epidemiology of quality. *Inquiry* 1985;22:282-292.
13. Kessner DM, Kalk CE, James S: Assessing health quality—the case for tracers. *N Engl J Med* 1973;288:189-194.
14. Rhee KJ, Donabedian A, Burney RE: Assessing the quality of care in a hospital emergency unit: A framework and its application. *Quality Rev Bull* 1987;13:4-16.
15. Williamson JW: Formulating priorities for quality assurance activity: Description of a method and its application. *JAMA* 1978;239:631-637.
16. New York Academy of Medicine, Committee on Public Health Relations: *Maternal Mortality in New York City; A Study of All Puerperal Deaths 1930-1932*. New York, Oxford University Press Inc, 1933.
17. Kohl SG: *Perinatal Mortality in New York City: Responsible Factors*. Cambridge, Mass, Harvard University Press, 1955.
18. Rutstein DB, Berenberg W, Chalmers TC, et al: Measuring quality of medical care: A clinical method. *N Engl J Med* 1976;294:582-588.
19. Mushlin AI, Appel FA: Testing an outcome-based quality assurance strategy in primary care. *Med Care* 1980;18:1-100.
20. Donabedian A: *The Criteria and Standards of Quality*, vol 2: *Explorations in Quality Assessment and Monitoring*. Ann Arbor, Mich, Health Administration Press, 1982.
21. Donabedian A: Criteria and standards for quality assessment and monitoring. *Quality Rev Bull* 1986;12:99-108.
22. Greenfield S, Lewis CE, Kaplan SH, et al: Peer review by criteria mapping: Criteria for diabetes mellitus: The use of decision-making in chart audit. *Ann Intern Med* 1975;83:761-770.
23. Greenfield S, Cretin S, Worthman L, et al: Comparison of a criteria map to a criteria list in quality-of-care assessment for patients with chest pain: The relation of each to outcome. *Med Care* 1981;19:255-272.
24. Lyons TF, Payne BC: The use of item weights in assessing physician performance with predetermined criteria indices. *Med Care* 1975;13:432-439.
25. Stewart AL, Ware JE Jr, Brook RH: Advances in the measurement of functional states: Construction of aggregate indexes. *Med Care* 1981;19:473-488.
26. Fanshel S, Bush JW: A health status index and its application to health service outcomes. *Operations Res* 1970;18:1021-1060.
27. Patrick DI, Bush JW, Chen MM: Methods for measuring levels of well-being for a health status index. *Health Serv Res* 1973;8:228-245.
28. Weinstein MC, Stason WB: Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977;296:716-721.
29. Willems JS, Sanders CR, Riddiough MA, et al: Cost-effectiveness of vaccination against pneumococcal pneumonia. *N Engl J Med* 1980;303:553-559.
30. Peterson OL, Andrews LP, Spain RA, et al: An analytical study of North Carolina general practice, 1953-1954. *J Med Educ* 1956;31:1-165.
31. *What Sort of Doctor? Assessing Quality of Care in General Practice*. London, Royal College of General Practitioners, 1985.