

Advantages and Limitations of Explicit Criteria for Assessing The Quality of Health Care

AVEDIS DONABEDIAN

*School of Public Health,
University of Michigan*

THE PRIMARY MOTIVATION FOR DEVELOPING explicit criteria and standards for quality assessment in health care has been the striving for consistent and valid judgments. Having reviewed the methods then available for quality assessment, and having found them wanting, Lembcke (1956, 1959) was impelled to develop a method that involved the rigorous use of explicit criteria so that it would merit being called “scientific.” The reasoning that led a Michigan group to a rediscovery of the “criteria approach” was essentially similar. Neither case-by-case review, as then practiced, nor the use of statistical norms to declare individual cases deviant, offered a satisfactory solution to the problem of making valid judgments of understay and overstay in the hospital. A new method was needed. In the words of the authors,

Review of cases in which all the facts for one case are considered at one time and balanced judgment attempted tends to result in subjective and lenient decisions. In spite of the values of the method, it encourages the concept of each case as a purely individual one; considerable thinking in the area of medical care is conditioned by this approach, which is essentially obscurantist in effect. The opposite error—treating patients as statistics or interchangeable units—is also essentially fruitless, because it cannot penetrate deeply enough to yield the necessary knowledge.... One way out of this

dilemma is to review individual cases but to apply general criteria of effectiveness to each (Fitzpatrick et al., 1962:454).

Whether the explicit criteria approach is, in fact, superior in reliability and validity to one that uses implicit criteria alone is, of course, subject to empirical verification.

Another reason for offering the explicit criteria approach is its greater simplicity and lower cost. The formulation of the criteria can be time-consuming, and requires a great deal of professional knowledge and other skills. But once the criteria have been selected, and the standards specified, a complicated and voluminous medical record can be reduced to a concise summary by abstracting the information called for by the criteria and standards. The abstract can be prepared, under supervision, by trained nonprofessionals, and a computer can be used to collate, arrange, and display the relevant information. In this way one reduces to a minimum the use of health care professionals whose time is exceedingly costly, and whose interest in the review process is generally less than enthusiastic. As Lembcke (1959:65) puts it,

It is said that with a cookbook, anyone who can read can cook. The same is true, and to about the same extent, of the medical audit using objective criteria; anyone who knows enough medical terminology to understand the definitions and the criteria can prepare the case abstracts and tables for the medical audit. However, the final acceptance, interpretation and application of the findings must be the responsibility of a physician or group of physicians.

The purer forms of implicit review have a voracious appetite for professional time. "Furthermore," as Morehead (1976:118) points out, "years of experience with this approach have made it clear that not all physicians, even the most eminent, can perform this task in a constructive, analytical fashion." Consequently, the successful use of the method requires great attention to the selection and training of the reviewers; and the process of review makes continuous demands on their knowledge, judgment, and attention. Of course, the selection of those who are to formulate explicit criteria is equally, if not more, critical, since the result of their decisions can often have a widespread effect. However, once the criteria have been made explicit, their reasonableness and validity can be directly verified. When a reviewer

of the quality of care begins by using implicit criteria, we must depend entirely on his judgment and integrity, unless he reveals, in detail, the reasons for his judgments.

The degree of explicitness and specification in the criteria may also be related to the ability to permeate and influence the conduct of care, so that it conforms to more general institutional or social objectives. Explicit criteria may be viewed as an instrument of control. By codifying a certain view of what is meant by good medical practice, they are capable of both reflecting and influencing social and institutional policy. He who controls the criteria controls a key element in the system, but only to the extent that the criteria can be made effectively operational in everyday practice. By contrast, implicit criteria, though they reflect the general norms of a profession, are by their very nature less amenable to large-scale programming and control. They represent more accurately the more local and individualistic traditions in the organization and practice of the professions.

Those who work in organized settings need to know by what standards they are to be judged, and to be assured that these are applied consistently and fairly. The explicit criteria are more likely to meet these expectations, provided the concept of quality embodied by the criteria is acceptable and complete. If not, strict limits are likely to be placed upon their use; and, in some instances, they may be ignored or actively opposed. Thus, the two forms of criteria are adapted to two requirements that seem contradictory, and yet simultaneously necessary to the proper control of professional behavior. The explicit criteria respond to the need for predictability, consistency, and fairness. The implicit criteria are needed to accommodate legitimate professional considerations that are not represented in any particular set of explicit criteria. For these reasons, in everyday practice the judgments based on explicit criteria are subject to review. That this review may also be more tolerant of the failings that beset everyday practice is something much sought after by the practitioner, though it may be deplored by others.

It has been claimed, with some reason, that the formulation of explicit criteria is a worthwhile enterprise in itself. It can open a discussion of the social and scientific bases of practice, leading to an exploration of both social legitimacy and scientific validity. The scope of the exploration and its consequences would, of course, depend on who participates. While admitting that, almost always, these dis-

cussions are dominated by the profession whose practice is being codified, one must recognize that there is an opportunity of representing a broader variety of views, including those of the consumer. Even when participation is restricted, the discussions, besides being highly educational to all the participants, should serve to verify the professional and social validity of the criteria, and to help bring about a consensus in their favor and a commitment to their implementation. And the implementation of uniform criteria should lead to greater fairness in access to more equal care—across programs, geographic locations, and institutional or individual providers.

The formulation of explicit criteria and standards of health care carries with it a greater specification of other related issues and phenomena. As a background, one needs a definition of the scope and objectives of care, at least in general terms, perhaps with greater specification of the meaning of quality and of the approaches to its assessment. The specification of process criteria may also require a prior specification of the level of quality to be attained, and of the nature of the system that provides care (Lee and Jones, 1933; Schonfeld et al., 1975). The specification of both process and outcome criteria is so intimately dependent upon the prior construction of homogeneous categories of patients to be cared for, that specificity in one cannot be achieved without corresponding specificity in the other (Brook et al., 1977; Williamson, 1978). Similarly, the lists of explicit criteria seem to cry out for specification of the way in which degrees of adherence to the criteria can be translated into a descriptive judgment of goodness in care.

The need to carefully select and define the conditions to which the explicit criteria pertain makes it impractical, if not impossible, to prepare criteria sufficiently specific for the large variety of conditions, and of their combinations, that constitute the totality of practice. Some method for selecting conditions is needed, raising complex issues about the basis for selection, and its consequences. By contrast, if implicit review is used, the entire range of conditions embraced by any of the recognized fields of practice can be included, so that a representative picture of that practice can be obtained.

The most important criticism of the explicit criteria approach is that it may achieve higher levels of reliability at the expense of reductions in validity. In the words of Morehead et al. (1964:41),

Frequently, such criteria force into a rigid framework similar actions or factors which may not be appropriate in a given situation due to the infinite variations in the reaction of the human body to illness.... The study group rejects the assumption that such criteria are necessary to evaluate the quality of medical care. It is their unanimous opinion that it is as important for the surveyors to have flexibility in the judgment of an individual case as it is for a competent physician when confronting a clinical problem in a given patient.

In a lifetime of work that spans the modern era of quality assessment, almost from its very beginnings, Morehead has remained faithful to this view, not for want of reexamination, but because experience has seemed to reaffirm its validity. That is not to say that she has been unaware of the need for structuring the process of judgment to the extent possible, so that it can be carried out systematically and skillfully, and its rationale exposed to scrutiny. In all her studies, a great deal of attention is paid to the selection of reviewers. In the second of her studies of hospital care for members of the Teamsters' Union, two judges reviewed each record independently, justified their conclusions in writing, and discussed their differences. In the many earlier studies of ambulatory care at the Health Insurance Plan of Greater New York (Daily and Morehead, 1956; Morehead, 1967), as well as in the more recent studies of neighborhood health centers and other clinics (Morehead, 1970; Morehead et al., 1971; Morehead and Donaldson, 1974), and hospitals (Fine and Morehead, 1971), the process of review has been guided by specification of the elements of care to be judged, and of the methods for arriving at numerical scores and descriptive judgments summarizing the quality of care.

Perhaps reflecting the more general trend, the degree of specification in Morehead's work has seemed to grow with time; and, whenever possible, she and her associates have not been averse to identifying a set of "basic" procedures and services that apply to an entire category of patients; see, for example, Morehead (1970) and Fine and Morehead (1971). Nevertheless, the ultimate reliance on the clinical judgment of an expert who reviews the entire record of care has remained unshakable. And in this, I must confess that I agree with her. For although peer review of the entire record of performance (whether of process alone, or of process and outcome combined) is open to error

and abuse, as we all recognize, there is nothing we now have that can handle better the entirety of practice in all its rich variety and detail. It is true that, in everyday use, this method can become, as Fitzpatrick et al. (1962) warn us, "obscurantist in effect." But the obscurantism of the implicit approach is a consequence of its misapplication; whereas the explicit criteria are open to an obscurantism that is incorporated into their essence and form, so that they are in danger of becoming instruments of institutionalized and pervasive error.

Of course, there is nothing nefarious about explicitness itself, unless one fears scrutiny and challenge. The faults of explicit criteria, when present, come either from imperfections in design or from misuse. But there is a sense in which some of the virtues of explicit criteria are also the progenitors of their failings. For example, in order to achieve specificity and a reasonable level of completeness, there is a temptation to attenuate the definition of quality until it is no more than a shadow of its more real, more fully rounded self. And, as expected, what survives is the "harder" core of technical concerns, while most subject to loss are the "softer" elements of interpersonal and social management that many professionals continue to view as more peripheral. But no matter how broadly quality is defined, to equate quality with a list of procedures necessary for, or consistent with, the care of a given diagnosis, is to accept a caricature that has lost all the finer shadings with which clinical judgment adorns the true face of excellence. The shorter the list, the more niggardly the standard of quality is likely to be; the longer the list, the greater the temptation for indiscriminate and wasteful use. In general, the proponents of explicit criteria have found it easier to specify clearly necessary care than to define what is precisely optimal, and what is redundant.

The greater amenability of explicit criteria to being used as an instrument of control is also a two-edged sword. In this capacity, their utility and their dangers stem not only from their design, but also, and more important, from who uses them, in what way, and for what purpose. Properly constructed and used, explicit criteria can expand the definition of quality and raise its level. Improperly used, they can impose an oppressive and misguided uniformity, assuming the professions allow themselves to be so dominated.

Notwithstanding their limitations, explicit criteria have no diffi-

culty at all in identifying, with commendable decisiveness, many of the grosser deficiencies of technical care that are, unfortunately, only too frequent today. In a more general sense, the usefulness of explicit criteria is reinforced, and their limitations mitigated, when the criteria are used, not as full representations of quality, but as screening devices, to separate care of doubtful quality from that which is likely to be acceptable. In this context, some degree of error, in either direction, is both inevitable and tolerable. Improvements in the design and application of explicit criteria can reduce this error still further. In this regard, the development of branching or algorithmic criteria, as Greenfield et al. (1975) have shown, is particularly promising. Eventually, further improvements in the explicit criteria may make them acceptable and reasonably complete representations of the quality of care. Until then, the very presence of the explicit criteria may be a temptation for the unwary to fall into that error of "misplaced concreteness" against which Alfred North Whitehead has warned.

References

- Brook, R.H., Davies-Avery, A., Greenfield, S., Harris, L.J., Tova, L., and Ware, J.E., Jr. 1977. Assessing the Quality of Medical Care Using Outcome Measures: An Overview of the Method. *Medical Care* 15 (Supplement to September 1977):1-165.
- Daily, E.F., and Morehead, M.A. 1956. A Method of Evaluating and Improving the Quality of Medical Care. *American Journal of Public Health* 46 (July):848-854.
- Fine, J., and Morehead, M.A. 1971. Study of Peer Review of In-hospital Patient Care. *New York State Journal of Medicine* 71 (August 15):1963-1973.
- Fitzpatrick, T.B., Riedel, D.C., and Payne, B.C. 1962. The Effectiveness of Hospital Use. In McNerney, W.J., and Study Staff, *Hospital and Medical Economics: Services, Costs, Methods of Payment, and Controls*, Vol. I, 449-455. Chicago: Hospital Research and Educational Trust.
- Greenfield, S., Lewis, C.E., Kaplan, S.H., and Davidson, M.B. 1975. Peer Review by Criteria Mapping: Criteria for Diabetes Mellitus: The Use of Decision-Making in Chart Audit. *Annals of Internal Medicine* 83 (December):761-770.
- Lee, R.I., and Jones, L.W. 1933. *The Fundamentals of Good Medical Care*. Chicago: University of Chicago Press.

- Lembcke, P.A. 1956. Medical Auditing by Scientific Methods: Illustrated by Major Female Pelvic Surgery. *New England Journal of Medicine* 162 (October 13):646-655.
- . 1959. A Scientific Method for Medical Auditing. *Hospitals* 33 (June 16):65-71, and 33 (July 1):65-72.
- Morehead, M.A. 1967. The Medical Audit as Operational Tool. *American Journal of Public Health* 57 (September):1643-1656.
- . 1970. Evaluating Quality of Medical Care in the Neighborhood Health Center Program of the Office of Economic Opportunity. *Medical Care* 8 (March-April):118-131.
- . 1976. P.S.R.O.: Problems and Possibilities. *Man and Medicine* 1 (Winter):113-123.
- , and Donaldson, R.S. 1974. Quality of Clinical Management of Disease in Comprehensive Neighborhood Health Centers. *Medical Care* 12 (April):301-315.
- , Donaldson, R.S., and Seravelli, M.R. 1971. Comparisons between OEO Neighborhood Health Centers and Other Health Care Providers of Ratings of the Quality of Health Care. *American Journal of Public Health* 61 (July):1294-1306.
- , Donaldson, R.S., Sanderson, S., and Burt, F.E. 1964. *A Study of the Quality of Hospital Care Secured by a Sample of Teamster Family Members in New York City*. New York: Columbia University, School of Administrative Medicine.
- Schonfeld, H.K., Heston, J.F., and Falk, I.S. 1975. *Standards for Good Medical Care*, Vol. I. Washington, D.C.: U.S. Government Printing Office. DHEW Publication No. (55A) 75-11926.
- Williamson, J.W. 1978. *Assessing and Improving Health Care Outcomes: The Health Accounting Approach to Quality Assurance*. Cambridge, Mass.: Ballinger.

Acknowledgments: This paper is an excerpt from a forthcoming book, *Explorations in Quality Assessment and Monitoring. Volume II: The Criteria and Standards of Quality* (Ann Arbor, Michigan: Health Administration Press), which is being prepared with joint support from the National Center for Health Services Research (under Grant No. HS-02081) and the Commonwealth Fund. The opinions expressed in the paper are those of the author. They do not in any way represent his sponsors.

Address correspondence to: Professor A. Donabedian, Nathan Sinai Distinguished Professor of Public Health, School of Public Health, The University of Michigan, 109 Observatory Road, Ann Arbor, MI 48109.