INVolVING ConSUMERS IN QuAlITY oF CaRE aSSESSMENT: . . .
D0 THEY PROVIDE VaLID INfoRMATION?

Allyson Ross Davies
John E. Ware, Jr.

December 1987
ACKNOWLEDGMENTS

The authors acknowledge the support of the Hartford Foundation during the preparation of this paper. This paper contains the same text as an article that will appear in the Spring 1988 issue of Health Affairs, Volume 7, No. 1, which because of space considerations included an abbreviated reference list. The reference list in this version is complete.
INTRODUCTION

Few would question that the delivery of quality medical services is a time-honored precept of health care providers. History indicates, however, that efforts to measure the quality of the care actually provided and to determine whether it meets standards come and go.\textsuperscript{1} Interest in quality of care assessment and assurance (QA) activities is again on the ascent, fueled in part by concerns that the cost-containment policies of the early 1980s were implemented with little attention to their effects on the quality of services (or on the health of patients). The pressures of a more competitive marketplace also serve to focus current attention on quality, as providers seek to improve their services and to distinguish them from those of others.

Along with this resurgence of interest in quality of care comes renewed attention to the consumer's viewpoint, and efforts to involve consumers in ongoing QA activities are likely to increase rapidly. Although several factors may motivate decisions to involve the consumer in such activities, one assumption appears to be critical: that consumers can provide valid information about the quality of their medical care.

In attempting to determine whether this assumption is correct, we outline the major arguments offered for and against using information provided by the consumer in QA activities, and review the evidence that supports or refutes these arguments. We aim to identify those QA activities that can rely on data from consumers because the evidence indicates they are valid for that purpose. Where empirical evidence is lacking, we identify the types of information needed to warrant use of
consumers' data by other QA activities. Before this review, we pause to define a term and make an important distinction.

"Quality of care" refers to the worth or excellence of various attributes of medical care. Most definitions focus on the encounter or episode and include such attributes as the \textit{technical process} of care (e.g., diagnosis and management) and \textit{interpersonal features} of the provider-patient relationship.\textsuperscript{2-5} Broader definitions, pertinent to evaluating delivery systems rather than individual providers, also include such attributes as accessibility and availability of providers and services. We focus here on the narrower definition, because most of the arguments offered against using data from consumers question how accurately those data reflect attributes of the encounter or episode, and because available evidence supports the validity of consumers' information about access to care.\textsuperscript{6-9}

It is important to distinguish data of two types: ratings and reports. Display 1 illustrates the difference between ratings and reports about two features of medical care quality (technical process and interpersonal aspects). Consumers' \textit{ratings} capture personal evaluations of attributes of medical care providers and services; they are inherently more subjective because they reflect both personal experiences and the standards consumers apply when evaluating care. By contrast, consumers' \textit{reports} deal with things that did or did not occur; they are inherently more objective and potentially can be confirmed by an outside observer. Conclusions regarding the usefulness of data from consumers in QA activities may well differ depending on whether reports or ratings are considered.
Display 1. Distinction between consumers' ratings and reports regarding technical and interpersonal aspects of care

<table>
<thead>
<tr>
<th>Technical</th>
<th>Interpersonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td></td>
</tr>
<tr>
<td>evaluation (excellent-poor) of doctor's</td>
<td>evaluation (excellent-poor) of courtesy and</td>
</tr>
<tr>
<td>thoroughness</td>
<td>respect shown by doctor</td>
</tr>
<tr>
<td>Report</td>
<td></td>
</tr>
<tr>
<td>indication (yes-no) of whether doctor checked</td>
<td>indication (yes-no) of whether doctor introduced</td>
</tr>
<tr>
<td>blood pressure</td>
<td>self</td>
</tr>
</tbody>
</table>

THE ARGUMENTS AGAINST

Arguments that call into question the validity of information provided by consumers about quality of care have one thing in common. They all presume that the information reflects something other than attributes of medical care. Thus, they argue that data from consumers:

- ...reveal more about the consumer than about the quality of care.

Consumers' ratings of medical care correlate with individuals' attitudes toward the community, satisfaction with life, and values or expectations regarding medical services; 7,10-14 with health status; 15,16 and with such educational level, age, income, ethnicity, and geographic location. 7,17-22 Because consumers' ratings of medical care contain information about the consumers, some question their validity as assessments of medical care. 13,16,23

This evidence in fact is hard to interpret, because we know little about the actual relationships among these factors, and thus about what should be the magnitude of the correlations we estimate. For example, do better educated consumers evaluate the interpersonal aspects of their
care less favorably because their providers treat them less personally, or because their judgments are biased? This first argument presupposes that ratings should reflect information only about attributes of medical care, and thus that any relationship between ratings and the rater's characteristics indicates bias. This argument fails to recognize, however, the real probability that features of medical care and individual characteristics are related in the real world, and that those with different characteristics receive medical care of differing quality. Moreover, this argument does not consider the weakness of observed relationships (e.g., on the order of 0.30 or less) and the inconsistent findings across studies regarding the relationships among such factors as age, education, and ethnicity and ratings of care. 24,25

Much of the evidence cited to support this argument comes from correlational and cross-sectional studies in which the actual attributes of medical care are unknown. Experimental studies that manipulate elements of the encounter or longitudinal studies that include an intervention provide more exacting validity tests, because they separate the effects of real differences in quality of care from bias related to self-selection. Observational studies that identify the elements of care rendered also provide useful information about validity. Several such studies have been done since Lebow noted their virtual absence in his 1974 review of consumers' assessments of quality. 23

Results from experiments by Ware et al. illustrate the true nature of the relationship between individual characteristics and attributes of care. 25,26 Subjects from a relatively young (median age 36 years) and well-educated (median 16 years education) sample were randomly
assigned to view one of four simulated physician-patient encounters. The videotaped encounters, each depicting a visit for acute pharyngitis, varied the elements of the history and physical (technical process) and of the physician's manner in relating to the patient (interpersonal aspects). Before viewing a tape, subjects provided sociodemographic data and rated their own medical care. After the viewing, they rated the technical and interpersonal aspects of care depicted in the simulation. While the pre-viewing ratings of their own care correlated significantly \((p < 0.01)\) with both age \((+0.26)\) and education \((-0.32)\), none of the eight post-viewing ratings was significantly related to age or education.\(^{25}\)

Also using a videotaped and scripted doctor-patient encounter to control the "stimulus," Linn et al. found that ratings of technical quality of the encounter were unaffected by importance assigned to four aspects of medical care (technical quality, psychosocial concerns, courtesy, and mutual participation). By contrast, ratings of interpersonal aspects were influenced by the importance assigned psychosocial concerns and technical quality.\(^{13}\)

A recent matched-groups observational study examined the effects of loss of health insurance on consumers' ratings of care. The reduction compromised both access to and quality of care for patients with chronic diseases (e.g., hypertension, diabetes). Mirroring these effects, Lurie et al. found that ratings by those who had lost coverage declined significantly relative to previous levels and to the control group.\(^{27,28}\)

Thus, results from several studies suggest that bias from personal characteristics is minimal for ratings of technical quality. Taken
together, the available evidence does not support the conclusion that this bias is strong enough to invalidate consumers' ratings of the quality of their care.

- **reflect how much was done, not how well it was done.**

The proponents of this viewpoint hold that consumers can be "seduced" by the kind or number of tests and procedures received into believing that the services provided were appropriate and well-performed. In so stating, they argue that more is not necessarily better. While quantity alone does not connote quality, the development and application of ever more medical technologies has contributed to a public impression that quantity is related to quality in the eyes of the medical profession. Thus, we might expect that quantity would influence consumers' ratings. At issue is whether that influence is so great as to mean that "quality" ratings are better interpreted as "quantity" reports.

Results from two experiments illustrate the influence of quantity on consumers' ratings of quality. Sox et al. studied middle-aged male patients with nonspecific chest pain who had been determined by clinical algorithm not to require tests to rule out myocardial infarction. Patients who were randomly assigned to receive two common but for them medically unnecessary tests (routine electrocardiogram, serum creatine phosphokinase) evaluated their overall care significantly more favorably than did those who did not get the tests.²⁹

In an experimental study of health insurance options, Davies et al. compared non-aged general population samples who were randomly assigned to a health maintenance organization (HMO) or to fee-for-service (FFS)
care. Despite the fact that the HMO enrollees consumed fewer services on average (particularly hospital services), no differences in average ratings of technical quality were observed. However, upper-income HMO enrollees who began their experimental participation in poor health were significantly less satisfied with technical quality than their FFS counterparts. As noted, the pattern of findings suggested that they may have equated less care with poorer quality care.\textsuperscript{9}

Because the populations and settings in these two studies differed, we should not be surprised that evidence regarding the influence of quantity on ratings of technical quality is not entirely consistent. The evidence does suggest that consumers' ratings of care do reflect, at least in part, how many services they received. Until the actual relationship between quantity and quality is better understood, and until these findings are replicated, the conclusion that this influence is so great as to invalidate consumers' ratings is unwarranted.

\begin{itemize}
\item \textit{...disagree with physician judgments regarding quality.}
\end{itemize}

Advocates of this viewpoint regard physician judgments as the criterion against which other evaluations of quality should be validated. They generally hold that consumers lack appropriate standards for judging the science and technology of medical care because they are medically untrained.\textsuperscript{30,31} Thus, this argument challenges consumers' ratings of technical rather than interpersonal aspects of care.

This argument encounters its first difficulty in holding consumers to a standard that itself is far from uniform. Evidence suggests that physicians do not always agree with their own or with each others' judgments of technical quality.\textsuperscript{32-34} Moreover, we would not expect
consumers' ratings to agree perfectly with those of physicians because the two groups have different perspectives regarding the process of care. 9,23

In an early study, Ehrlich et al. examined the congruence between consumer and physician judgments of quality of care. While consumers tended to rate care more favorably than did physicians, there did appear to be better agreement in judgments of poor care, the focal point of most quality assurance efforts. 35

Results from two sets of experiments provide evidence that strongly suggests consumers can evaluate technical quality of care for common problems, and that their ratings agree with those of physicians. The Ware et al. experiments, done first in a relatively young and well-educated sample, were replicated in an elderly (median age 72 years) and less educated sample (median 12 years education). 25

Using a similar approach, Chang et al. videotaped six simulated provider-patient encounters that deliberately manipulated three aspects of care (technical, psychosocial, and patient participation); courtesy (an attribute of interpersonal quality) was held constant. All tapes depicted a return office visit for previously diagnosed angina. Elderly ambulatory women (mean age 70.9 years) were randomly assigned to view one of the videotapes and then rated different features of the simulated encounter. 17

The manipulations of different attributes of the care in both sets of experiments were verified by providers. 17,25,26 Participants in both rated the technical process of the "high" technical quality encounters (scripted to include necessary and sufficient history and
physical examination items) significantly more favorably than the "low" technical quality encounters (which omitted relevant and included irrelevant items). To illustrate, Display 2 presents mean technical quality ratings by consumers for each of four simulated encounters in Ware et al. Under both interpersonal conditions, ratings of technical quality were significantly less favorable (p < 0.0001) for "low" than "high" technical quality encounters.

Display 2. Mean ratings of technical quality of care by consumers viewing one of four simulated encounters

<table>
<thead>
<tr>
<th>Interpersonal Aspects</th>
<th>Technical Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>High</td>
<td>18.00</td>
</tr>
<tr>
<td></td>
<td>(4.02)</td>
</tr>
<tr>
<td>Low</td>
<td>16.09</td>
</tr>
<tr>
<td></td>
<td>(4.15)</td>
</tr>
</tbody>
</table>

NOTE: Scores on the four-item technical quality scale could range from 4 to 20; higher scores indicate greater satisfaction. Standard deviations in parentheses; N=109.

This evidence indicates, at least for common problems, that consumers can distinguish between the technical aspects of care judged good and less-than-good by physicians.

- ...simply reflect whether the provider was nice to them.

Reports and ratings of the interpersonal aspects of the provider-patient relationship should reflect "whether the provider was nice." At issue here is whether the provider's interpersonal manner interferes with the patient's accurate assessment of the technical
process of care.

Apparently supporting this argument are the relatively high correlations (on the order of 0.60 to 0.70) observed between consumers' ratings of interpersonal and technical quality of care.\textsuperscript{7,24,36} These relationships, however, must also be evaluated against what we know of the "real world." Only if technical process and interpersonal manner are not closely related in physicians' practices would such correlations be evidence of bias in consumers' ratings.

The most telling evidence comes from experiments that deliberately altered the relationship between technical process and interpersonal aspects of care, and from observational studies that carefully codified elements of the physician's behavior. Analyses of the experiments described earlier tested the hypothesis that ratings of "high" and "low" technical quality encounters would not differ when coupled with the "high" interpersonal manipulation; neither study observed such an effect.\textsuperscript{17,25,26} As illustrated in Display 2 with data from the Ware et al. study, the manipulation of interpersonal aspects had an insignificant effect on technical quality ratings ($F=2.29$, $df=1,105$, $p=0.13$).\textsuperscript{25}

Supporting results have been reported by several observational studies, and indicate that ratings of interpersonal aspects reflect documented variations in those features. Korsch et al. coded the interpersonal aspects of pediatric encounters, and found they were significantly related to maternal ratings of interpersonal aspects of care.\textsuperscript{37} Stiles et al. observed and quantified distinct elements of providers' manner during encounters with patients, and reported it was
significantly linked to ratings of interpersonal aspects of care obtained after the encounter. DiMatteo et al. used objective methods to document the verbal and nonverbal affective behavior of resident physicians when treating their patients. Patients' ratings of interpersonal aspects of care distinguished between residents who differed in such behavior; their ratings of technical process did not so discriminate.

While the evidence is persuasive, certain limitations must be noted. Results from both experimental and observational studies relate to ratings of outpatient care. Our ongoing literature review has identified only one study that examined the consumer's ability to rate inpatient care, and none of their ability to rate technical process of inpatient care uninfluenced by the interpersonal aspects of that care. Moreover, available results pertain to relatively common situations (i.e., an encounter for a sore throat; management of angina in an elderly patient). We know little about whether interpersonal features of care influence ratings of technical process for more complex problems or episodes of care. Finally, some may object that results from experiments with simulated encounters will not generalize because the "consumer" who rated the care did not experience that care as a patient. Anecdotal evidence suggests that the subjects in these experiments did become "involved" in the simulations; the accuracy of recall for elements of simulated encounters further supports the claim that subjects do "experience" the simulated encounters. Results from observational studies that quantify both the technical and interpersonal features of care and obtain ratings of each are required
to answer this objection directly. Thus, the weight of the available
evidence suggests that interpersonal features of the encounter do not
obscure consumers' ability to distinguish levels of technical quality
that are judged different by providers.

...AND THE ARGUMENTS FOR

- Consumers' assessments of quality predict their behavior in the
  marketplace.

The current marketing strategies of prepaid and other group plans,
health insurance benefits, and hospital services appear to presume the
importance of consumers' views of such quality-related features as
provider continuity, comprehensive services, and good health outcomes to
their choices among health care alternatives. The literature indicates
considerable support for this presumption. Behaviors linked to
consumers' ratings of quality-related features of care include
doctor-shopping, complaints and disenrollment, compliance, recommending
a source of care to others, and utilization.

Consumers who hold more favorable attitudes toward the technical and
interpersonal features of their care are significantly less likely to
change physicians. Studies of attitudes toward care received
from prepaid health plans in California revealed that Medicaid
enrollees' ratings of both technical and interpersonal aspects of care
were significantly related (r = -0.86 and -.90, respectively) to
voluntary disenrollments. Plans with the lowest ratings had ten-fold
higher annual disenrollment rates (approximately 30 vs. 3 percent). Results from the same study revealed a significant positive relationship
between Medicaid enrollees' ratings of technical quality and whether
they would recommend their prepaid plan to others. More favorable ratings of care have also been linked to various measures of compliance in several studies. Finally, favorable ratings of care
have been linked to a greater likelihood that people will use health services.

Interestingly, the validity of consumers' assessments in relation to actual quality matters little to the proponents of this argument. Whatever "quality" means to consumers, the demonstrated relationship between their ratings of quality and health-related behavior justifies for many the collection of those ratings and programmatic attention to them.

- Consumers need not necessarily be qualified judges to provide accurate information about quality.

This argument recognizes the important distinction between ratings and reports. Most of the foregoing arguments against using consumer data in QA voice skepticism about the validity of consumers' ratings (or judgments). The evidence discussed earlier may serve to temper that skepticism, but it sheds little light on the accuracy of consumers' reports.

Whether consumers can accurately report what happens during the technical process of care is best judged in the context of known manipulations of that process. In addition to rating technical process and interpersonal care, participants in the Ware et al. experiments completed a process-of-care checklist after viewing the simulated encounter. They were asked to report whether history and physical items had or had not been performed.
Results confirm the accuracy of consumers' reports. First, overall accuracy tended to be very high, ranging from 80 to 94 percent (mean 87.5 percent) across the four groups in the younger sample. The elderly and less educated sample was somewhat less accurate (mean 70.5 percent, range 57 to 76 percent). Second, consumers' reports regarding medical history items were slightly more accurate (some 3 to 10 percent) than those regarding physical examination procedures. Third, consumers in the better educated sample were about as accurate overall in reporting what was and was not done as were the physicians who viewed the same tapes. Finally, quality-of-care algorithms developed by physicians and scored from consumers' reports distinguished "high" and "low" technical quality manipulations more accurately (94 percent correct classification) than did consumers' ratings of the technical quality (73 percent correct). 25,26

In examining four alternate methods for collecting data to assess physician performance in treating chronic obstructive pulmonary disease (COPD) against explicit criteria, Gerbert and Hargreaves compared the agreement between the methods (physician interview, patient interview, chart audit, and videotaped observation) in reporting physician behaviors that did and did not occur. (The patients had a mean age of 60.7 years, and been treated for COPD an average of 15.8 years.) With respect to tests ordered, patients and physicians agreed 96 percent of the time; for treatments mentioned, 94 percent; and on occurrence of patient education, 88 percent. 54

Such results suggest that consumers' reports hold considerable promise as a source of data for QA activities. Further research should
examine the extent to which accuracy is affected by such factors as whether the rater is an observer or the patient, the location and reason for the encounter, and sociodemographic characteristics.

- Data from patients are less expensive than other QA data sources.

This very practical argument derives from the belief that most common methods of QA carry high dollar and time costs. For example, medical record abstraction according to protocol requires professional input to development of algorithms, often relies on highly-trained personnel to do the abstraction, and depends on the ready availability of the record. The search for data sources that may cost less leads some to consider surveying consumers about their care.

Few published studies include any information about the relative costs of medical record audit, computerized claims analysis, or other methods of quality assessment. Anecdotal evidence suggests that acquiring information from consumers probably does cost less than that obtained from records.

For example, in fall 1987 HCFA awarded contracts totalling $300 million to Professional Review Organizations (PROs) to assess quality of care for 25 percent of the anticipated 22 million Medicare hospital admissions during the next two years. Thus, the cost of each review will average about $55.00; estimates suggest that the time required to do the actual review averages one-half hour.\textsuperscript{55}

Recent studies that abstracted medical records from national samples of physicians' offices and hospitals provide similar estimates. Abstraction took from one-half to one hour per record, and costs for the review and data entry ranged from about $15 to $22 per record.
Including training, supervision, and validation in this estimate, as they are in the PRO figure, the cost per abstract averaged some $45. 56

Despite the considerable literature on survey methods, surprisingly little has been published regarding the costs of consumer surveys. As Fowler notes, costs will vary markedly depending on such factors as questionnaire length, administration method, dispersion of the sample, availability and interest of potential respondents, and callback procedures. Although he does not quote costs, Fowler's evidence suggests that mail and telephone survey procedures cost markedly less than personal interviews. 57 Anecdotal evidence suggests that telephone surveys cost approximately $1 per minute on "average." A recent estimate came in at $45 per respondent for a half-hour questionnaire using a combination of telephone and mail techniques in a disadvantaged population.

In a noteworthy exception to this dearth of information, Gerbert and Hargreaves reported costs in their comparison of data on quality of outpatient care for COPD patients. Chart audit took approximately 15 minutes; interviews, about 20; and videotaped observations, 40 minutes (including the encounter itself). Interviews and chart audit cost approximately $35 each (exclusive of respondent payments); at approximately $42, use of videotapes cost about 20 percent more. 54

Obviously, we have little hard evidence on which to base any conclusions regarding the costs of patient ratings or reports relative to more traditional sources of QA data. Although the anecdotal evidence cited was readily available rather than representative, it does suggest that obtaining QA data from consumers is no more expensive and, under
many circumstances, is probably less expensive than data obtained from more traditional sources. Publication of data collection costs and more comparisons of the costs of data collection methods would greatly inform the choice of QA data sources.

- **Data from consumers provides information not otherwise available.**

  Until recently, most methods for evaluating quality of care focused on inpatient care, and relied solely on examination of the medical record or data based on the record. Along with the broadening scope of mandatory QA efforts (e.g., Medicare, Joint Commission on the Accreditation of Hospitals) beyond inpatient care has come recognition that traditional QA databases pose difficulties to review of outpatient care. Records are often incomplete, particularly with respect to recording negative findings, and they rarely capture information about the interpersonal aspects of care. Claims capture little detail about outpatient services, and access to medical records of the office-based physicians who provide most outpatient care is problematic. Although methods have been developed to capture data that describe the encounter, they focus largely on the volume and nature (but not the content) of the communication between doctor and patient, and are both cumbersome and expensive to use.

  Thus, the search for alternative sources of data for QA requires information about interpersonal aspects of care in particular and about outpatient care more generally. Certainly the consumer is best qualified to rate interpersonal aspects of care, and we have little evidence that questions the validity of consumers' ratings of interpersonal features. Moreover, virtually all the evidence we have
reviewed that supports the validity of consumers' ratings and reports of technical aspects of care comes from studies in outpatient settings. Such evidence suggests that consumers can provide at least some of the information not currently available from traditional QA data sources.

We have noted that data from consumers costs no more, and probably less, than data obtained from more traditional QA data sources. Similarity of cost, however, does not indicate similarity of content. With one exception, 54 we found no analyses of the information in records or claims that can be accurately duplicated by reports from consumers or their proxies. Until we can better identify which source provides the most accurate (and least expensive) information regarding different aspects of the technical process, we should regard data from consumers as a complement to those from other sources rather than as a substitute.

IN CONCLUSION

Considering the available evidence, those who have argued against using data from consumers in quality assessment and assurance activities appear to have somewhat overstated their case. Our review leads us to the following conclusions regarding each of the arguments preferred:

- Bias from personal characteristics is not so strong as to invalidate consumers' ratings of the interpersonal or technical quality of their care.
- Consumers' ratings of technical quality do reflect, at least in part, how many services they received.
- For common problems, consumers can distinguish between the technical aspects of care judged good and less-than-good by physicians.
- 19 -

- Interpersonal features of care do not obscure consumers' ability to distinguish levels of technical process for common outpatient problems.

- Whatever "quality" means to consumers, their perceptions of quality affect their choices among health care alternatives.

- Consumers' reports (as distinct from ratings) hold considerable promise as a data source for quality assessment and assurance activities.

- The costs of obtaining data from consumers are not higher, and are probably lower, than those for data from more traditional sources such as record audits.

- Consumers are the best source of data on the interpersonal aspects of care; moreover, consumers can provide some of the data on technical quality of outpatient care not available from such traditional data sources as claims and records.

No argument directly questioned the validity of consumers' ratings of interpersonal aspects of care, suggesting that most accept the consumer's viewpoint on such aspects as the criterion. That fact, combined with evidence indicating their ratings accurately distinguish known differences in interpersonal features of care (93 percent correct classification of "high" and "low" interpersonal in the Ware et al. experiments \(^{25,26}\), leads us to conclude that consumers' ratings of interpersonal aspects of care provide useful and valid information for quality assessment and assurance programs.

Our review of the evidence makes obvious the small number of systematic investigations of the validity of consumers' ratings. Consumers' reports have received even less attention. In reaching these conclusions, we put greatest weight on the well-controlled observational and experimental studies that were designed to address validity issues. Much of the available evidence, however, comes from correlational studies designed to address other issues. Although such studies yield
valuable validity information as a by-product, we can put less weight on their results.

While Lebow's request in 1974 for more studies designed to examine the validity of consumers' assessments of quality has been met, still more work is needed to address the arguments that remain only partially answered. Given the importance of generalizing findings, observational studies that use objective methods to document different features of medical care encounters will be critical to examining the consumer's ability to distinguish levels of technical process and to rate that process uninfluenced by interpersonal features of care for a wide range of medical problems and in both outpatient and inpatient settings. Such studies should also provide information about the relationship, in actuality and in consumers' ratings, between quantity of services and technical quality of care. Given the promise of consumers' reports for filling noteworthy gaps in available data on quality, particularly of outpatient care, the factors that affect accuracy of reporting deserve considerable attention. Finally, documentation and comparison of the costs and content of different QA data sources will help identify instances in which they yield the same data but one is more cost-effective than another, as well as those in which they complement each other.

Several studies currently under way should provide results relevant to these issues. Much of the current research on the doctor-patient relationship relies on observation; some uses simulations or experimental interventions. Virtually all studies in this area obtain ratings (and sometimes reports) of different features of care from the patient or
observer. Broader based, in that the research is an integral part of ongoing quality assessment and assurance activities, are at least two important studies of which we are aware. In developing techniques to evaluate residents in training programs, the American Board of Internal Medicine has focused its attention on consumers' ratings of interpersonal aspects (in particular, humanistic qualities and communication skills) in the outpatient setting.\textsuperscript{5,67} Focusing on a different arena and with a broader definition of quality, the Hospital Corporation of America and Harvard Community Health Plan have undertaken a joint effort with us to test the usefulness to their QA programs of consumers' ratings regarding all aspects of inpatient care.\textsuperscript{68} Both projects have been explicitly designed to yield information about the validity of consumers' ratings and reports.

In summary, the evidence indicates that data from consumers warrants use by those interested in assessing or assuring the quality of medical care. Indeed, data from consumers undoubtedly represent the centerpiece of any effort to examine the interpersonal component of quality. As regards the technical process of care, reports and ratings from consumers will likely prove a valuable supplement to data from more traditional sources such as the medical record.

Writing two decades ago during an earlier peak of interest in using consumers' data in quality assessment, Donabedian said that "outcomes [health status, consumers' attitudes toward care] remain the ultimate validator of the effectiveness and quality of medical care."\textsuperscript{2} That view was echoed recently by Secretary of Health and Human Services Bowen, who stated: "Any system that truly aspires to high quality must
accommodate personal preferences ... we should put more, not less, emphasis on what patients expect as the outcome of their treatment."

Thus, philosopher and policymaker agree that consumers should be involved in assessing quality of care. We conclude that the available evidence suggests consumers can provide valid data to quality assessment and assurance activities, particularly those focusing on interpersonal aspects of outpatient care.
REFERENCES


25. Ware, J.E., R.L. Kane, A.R. Davies, and R.H. Brook. The patient's role in assessing the quality of the medical care process, forthcoming.


56. Kosecoff, J. Fink and Kosecoff, Santa Monica, CA, personal communication, August 1987.


67. Swanson, D., American Board of Internal Medicine, Philadelphia, PA, personal communication, August 1986.

68. Batalden, P., Vice-President for Medical Care, Hospital Corporation of America, Nashville, TN, personal communication, April 1986.
