The evaluation of health care and efforts to maintain and improve quality in health care have very largely drawn on quantitative methods. Quantification has made possible precise expression of the extent to which interventions are efficient, effective, or appropriate and has allowed the use of statistical techniques to assess the significance of findings. For many questions, however, quantitative methods may be neither feasible nor desirable. Qualitative methods may be more appropriate when investigators are "opening up" a new field of study or are primarily concerned to identify and conceptualise salient issues. Various qualitative methods have been developed which potentially have an enormous role in assessing health care. This paper examines some of the more important forms that have been used in that assessment and outlines principles of good practice in the application of qualitative methodology. It is intended to encourage a wider use of qualitative methods in assessing health care and a greater appreciation of how much such methods have to offer.

The term "qualitative" is sometimes used quite loosely. We will review some of the methods to which the term is properly applied, but first what is not a qualitative study should be emphasised. Research based on a small number of patients or respondents should not be considered qualitative just because the sample size is too small to assess statistical significance. This is more likely to prove to be an inadequate quantitative study. Similarly, a study is not qualitative because it is based on answers to a questionnaire about subjective matters nor because data are collected by personal interview. If such data are analysed and reported largely in terms of frequencies and proportions of respondents expressing particular views, that is also a quantitative study. Qualitative research depends upon not numerical but *conceptual* analysis and presentation. It is used where it is important to understand the meaning and interpretation of human social arrangements such as hospitals, clinics, forms of management, or decision making. Qualitative methods are intended to convey to policy makers the experiences of individuals, groups, and organisations who may be affected by policies.

### Qualitative methods of data collection

**Qualitative methods of data collection**

Qualitative methods of collecting data encompass a range of approaches (box).

#### IN DEPTH INTERVIEWS

Perhaps the main method of obtaining data for qualitative analysis is by interview. Interviews are a particularly flexible method of gathering data, allowing the investigator to respond to the individual way in which respondents interpret and answer questions. There are various interview formats, but qualitative analysis requires "in depth" interviews so the interviewer can obtain more detailed information than is possible, for example, from an interviewer administered questionnaire, particularly regarding the perceptions and reasons behind respondents' statements. Interviews may be "semistructured," where the interviewer has a fixed set of topics to discuss, or "unstructured," where the interviewer has only very broad objectives in relation to the interview and is largely led by the respondent’s priorities and concerns. Unstructured interviews are most appropriate when subject matter is particularly complex or when investigators want to understand reasons for views or are exploring areas which have not previously been extensively described. An interview format increasingly used in health services research is the "critical incident technique," in which respondents are asked to recall the details of a particular experience such as hospitalisation.¹

Because in depth interviews require the active participation and judgement of the interviewer it is important that they be conducted by interviewers who have been appropriately trained. There is a range of general interviewing skills that are relevant to performance, such as an ability to demonstrate interest in the respondent without excessive involvement that may result in bias and an ability to ask both open ended facilitating questions and specific probes when relevant. The interviewer also needs to understand the general objectives of a study as well as the intentions behind specific questions. Another essential requirement is that interviewers be perceived by respondents as neutral with regard to the subject matter of an interview. For example, hospital staff are inappropriate to conduct interviews about patient satisfaction, and individuals perceived as judgemental are inappropriate for more sensitive topics.

The analysis of in depth interviews normally entails some degree of formal content analysis of what the respondent has said. For this reason investigators usually tape record interviews which can then be transcribed for...
detailed content analysis. Broad principles of content analysis for qualitative research are discussed below as, to some extent, they are common to all of the methods discussed in this paper.

An illustration of the use of in depth interviews is a study of the neurological management of chronic headache. Neurologists were unsure about the appropriateness of their role in this area. The initial purpose of the study was to identify the expectations of patients with chronic headache regarding their neurological referral, and to establish whether those expectations were satisfied. When interviewed before their outpatient appointments, patients proved quite unsure what to expect. Content analysis of these interviews instead disclosed a small number of different concerns felt by patients about their headaches – for example, a need to be reassured about serious disease or a desire to receive advice about how to change lifestyle or diet to avoid headaches. After their neurological consultations a minority of patients did experience serious disappointment with their clinic attendance but for various reasons were very uncomfortable with the language of “satisfaction” or “dissatisfaction” to describe their experiences. Overall the study suggested particular failures of communication in clinics that led to some patients’ concerns not being fully addressed.

In this example the primary advantage of in depth interviews was that they allowed the investigator to focus maximally on patients’ perceptions of their health problems and responses to health care rather than imposing his own categories. This led to a new and more meaningful typology of patients’ concerns. Examples of other qualitative analyses of in depth interviews include studies of dimensions of patients’ experiences of care for various chronic health problems, of the role of clinical audit in British medicine, of informal reasons for clinicians’ use of echocardiography, and of the concerns of singlehanded GPs in an inner London area.

FOCUS GROUPS
Groups rather than individuals may be interviewed by means of a technique commonly referred to as a focus group. Typically, eight to ten individuals are recruited to a group discussion about specified topics. The discussion is led by a trained moderator or facilitator and normally lasts for one and a half to two hours. The discussion is tape recorded, transcribed, and subjected to content analysis. Rather like the interviewer conducting in depth interviews, the moderator has a clear agenda of issues on which he or she stimulates discussion but aims to be fairly non-directive in encouraging discussion. Membership of a focus group needs to be reasonably homogeneous; too much heterogeneity in terms of social background or perspective on a topic tends to result in participants feeling inhibited from revealing views. Therefore if the purpose is to identify the range of views of individuals of different social backgrounds the normal practice is to conduct a series of focus groups.

Focus groups are particularly useful where investigators wish to establish quickly the range of perspectives on an issue of importance among different groups – for example, when it is felt necessary to “bridge a gulf” in understanding between providers of a service and the intended users. The dynamics of a well conducted focus group are such that individuals’ revelations of their views can “spark off” other participants to reveal broader insights than are possible from individual interviews. However, focus group methods have been compared to “pulling teeth” to stimulate discussion in a group that has no experience or interest in a topic. It is less appropriate in a focus group to probe for elaboration of individuals’ statements than is possible in an in depth interview, so that focus groups may be said to be stronger on breadth than depth. The facilitator is less in control of a focus group than is an interviewer with single respondents so it is a method that maximises the expression of perspectives not imposed by the researchers. There are some topics that are sufficiently sensitive that group discussion is unhelpful whereas individual interviews provide a more confidential context for self-revelation.

Focus groups are commonly regarded as an exploratory method, so that investigators may conduct them in order to design the questionnaire for a more definitive quantitative survey. Some would argue, however, that focus groups can provide valuable evidence in their own right, provided that investigators are concerned with conceptual rather than numerical analysis.

There are questions still to be addressed with focus group methodology. For example little is known about how the effects of social desirability and conformity influence expression of views and to what extent heterogeneity of participants influences results. Similarly, though facilitators clearly require substantial interpersonal skills such as ability to listen and to facilitate without becoming so involved as to bias discussion, facilitator effects on the quality of focus groups are not well understood. However, health is a particularly appropriate application for this method, which is being increasingly advocated for use by purchasers in obtaining local views about health and health care.

NOMINAL GROUP TECHNIQUES
For some purposes a more impersonal and less threatening form of group dynamics may be needed. A somewhat more structured form of gathering data from groups has been developed, called “nominal groups” because exchange and interaction between group members is more controlled than in focus groups. It is considered a technique less prone to the bias arising from vocal individuals influencing a group’s views in open discussion.

A nominal group normally is composed of eight members who meet together with a
Qualitative methods for assessing health care

leader who introduces the group’s tasks. The leader explains the question or problem on which the group’s views are sought. Individuals are asked to list on a paper form their different feelings or experiences in relation to the question (for example, different kinds of disappointments experienced by a group of users of a clinic), without discussion with other group members. Participants are then asked to declare their written comments which are recorded, as closely as possible to the participants’ own words, on a blackboard or flipchart. The complete list is then discussed by the group and a preliminary ranking made by the group from most to least important item of the total list. After discussion of this preliminary ranking a second, private ranking of items is performed by individuals on paper. This information constitutes the researcher’s core data. As with focus groups, investigators usually set up several different groups that are needed to represent the different perspectives with interests involved with an issue.

The primary advantages of this form of group research are that individuals may give more carefully considered expressions of their views compared with focus groups and are constrained by the tasks to produce more structured and explicit priorities. The structured nature of data gathering means that this is a method particularly appropriate to contexts where technical or complex issues need to be assessed. The primary disadvantage is that individuals are discouraged from “sparking off” each other, so that it is possible that the full range of views and observations are not elicited. Overall, like focus groups, nominal group techniques are a relatively low cost and quick method of exploring the parameters of an issue.

A variant of this methodology has been developed for use in developing professional consensus about appropriate indications for health care interventions. Panels representing relevant clinical expertise are provided with patient vignettes with varying symptoms and other indications. Vignettes are privately rated in terms of degree of appropriateness for the intervention. The group is then actually assembled, it discusses the ratings, and it is given the opportunity to revise prior decisions.

A recent example indicates that substantial agreement could be obtained regarding appropriateness for prostatectomy. Most of the steps in data gathering and interpretation of this application of nominal group techniques are actually quantitative, and readers are referred elsewhere for methodological problems associated with such techniques.

OBSERVATIONAL STUDIES

When the research question concerns what actually happens in health care settings, rather than participants’ perceptions of and responses to it, a more appropriate method of data collection may be direct observation. For this method the investigators attend the events they are concerned with, pay close attention to what goes on, and make a careful record of it for future analysis, using analytical techniques similar to those for other qualitative data.

Careful recording of what occurs is the cornerstone of observational research. Field notes, taken during or immediately after periods of observation, have been the traditional method of recording. The value of field notes depends on the skill and discipline of the investigator in recording as much detail as accurately as possible. Field notes are inevitably selective and so may be subject to observer bias. However, they are a relatively efficient way of recording observations, particularly in sensitive situations, and provide data in a form that is immediately accessible for analysis. Audio recordings and video recordings provide a more detailed and accurate record of events, although the demands of the technology may limit the range of events that can be observed and the tapes themselves present considerable additional demands in terms of transcription before analysis. Tape recordings are attractive in that they allow the investigator to go back and “reobserve” events, which may allow more detailed analysis than is possible with field notes. They also afford a record of the evidence on which interpretations are made which may be inspected by independent observers in establishing the validity of the analysis. However, audio recorders and particularly video recorders have been criticised as being so obtrusive as to change the nature of the events they are recording, thereby invalidating the observation itself. With smaller machines and consulting rooms specifically adapted for recording this is becoming less of a problem, and, in any case, those who have been the subject of audio or video recording have almost always found that the need to deal with the immediate demands to the situation constrain them to carry on in such the same way as they would if they were not being observed.

Investigators may collect data as “participant” or as “non-participant observers.” Participant observers participate in the daily life of the organisation over an extended period of time, watching what happens, listening to what is said, and asking questions. It is an exceptionally demanding way of gathering data. The researcher may encounter difficulties in being accepted by the group initially and then in sustaining the role long enough to observe the full range of events. The justification for such efforts in is the way participant observation enables the researcher to see and experience the institutional culture from the point of view of an “insider.” Goffman’s classic study of asylums illustrates well the unique insights that can be gained by this method of data collection. Goffman joined an American mental hospital as the assistant to the athletic director in order to observe the routines of the patients without drawing attention to himself. The resulting study develops the concept of the “total institution,” identifying its core elements and using it as a framework to describe and make sense of the culture of both the patients and
the staff. By closely examining what actually went on in the hospital, he was able to get behind official accounts of what the institution did to show the process and strategies that made it work in practice, and the consequences they had for patients' lives.

Participant observation has become less common as a research method, partly because of questions raised about the ethics of covert observation. Non-participant observation allows the researcher to remain as an accepted outsider, watching and recording the interactions as a "fly on the wall." Non-participant observation is particularly useful when the researcher is concerned to describe and conceptualise the "taken for granted" practices of everyday medical life: the routines and strategies that those they are studying develop in carrying out their work which may be so common and familiar as to be outside their conscious awareness. An example is Strong's study described in The Ceremonial Order of the Clinic in which he set out to elucidate the rules which govern interactions in medical consultations. On the basis of observations of 1120 paediatric consultations in several medical settings in Scotland and the United States he documented in detail the existence, nature, and scope of the rules which make up the limited number of institutionalised roles universally adopted by patients and doctors.

The main advantage of observation as a method of data collection is in allowing the investigators to "see for themselves," thus avoiding the biases inherent in participants' reports, such as selective perception, poor recall, and the desire to present themselves well. It is particularly appropriate when the investigator is concerned with describing and conceptualising how health services operate - that is, with practices and processes which are directly accessible to an outside observer. Its main drawback is that it limits the number and range of situations that can be studied to those at which the investigator is present. It is the most labour intensive form of data collection and perhaps for this reason is more often used in conjunction with other methods rather than as the main source of data.

CASE STUDIES
Case studies are not really a distinctive method of data collection or analysis but warrant separate discussion because they are increasingly used in the study of health care systems. They provide interpretations and analyses that are ultimately qualitative in nature. The most obvious respect in which case studies are qualitative is that typically they involve the study of a single or small number of units where quantitative manipulation of variables associated with units would be inappropriate. However, to a greater extent than the methods discussed so far, case studies often draw on mixture of quantitative and qualitative data. An example illustrates the point. In 1986 a resource management (RM) initiative was announced for the NHS, the purpose of which was to improve patient care by providing systems whereby managers and clinicians made better informed and more effective use of resources. The initiative was to be piloted in six acute hospitals and the costs and benefits were to be evaluated by a research team. Much quantitative data were gathered by investigators in relation to hospital activities and their costs in the pilot sites over time. However, as a number of other institutional changes to the NHS were simultaneously having effects in both the study hospitals and any potential "control" hospitals outside the pilot it was decided that it would be impossible to isolate quantitative benefits due to the resource management initiative specifically. Moreover, as with most human organisations, NHS defined objectives of the resource management initiative drifted over time. The authors concluded that qualitative evaluation of the initiative base on qualitative and quantitative evidence available was more informative. Qualitative inferences were drawn from the study such as that it was possible to involve clinicians in hospital resource management but that it was not clear that the high administrative and other costs associated with such involvement were matched by benefits to patient care.

Case studies such as the resource management study usually entail a combination of methods. Investigators carry out interviews, conduct participant observation on relevant meetings, and inspect written documents such as minutes or records. A systematic approach to interviews in an organisation can be adopted. For example, in a study of the impact of general management in the NHS, investigators, if confronted with discrepant accounts between different actors, would probe in interviews to obtain possible explanations. Case studies attempt to get an accurate picture by means of "triangulation," whereby the degree of convergence between different sources (for example, interview and documents) is carefully considered. A particular form of checking the plausibility of investigators' explanations is that of "respondent validation" in which the analysis that has emerged of a setting is presented to participants for their reactions. By means of this technique analysis can be refined and improved by respondents' feedback.

Qualitative analysis of data
If the distinguishing feature of quantitative evidence is the manipulation of numerical data then qualitative analysis is characterised by the development and manipulation of concepts. The investigator's primary tasks are the inspection and coding of his or her data in terms of concepts and then the manipulation of such concepts into analyses of underlying patterns. Although there are significant
differences of form or emphasis, this crucial role of conceptual analysis is common to all of the qualitative methods outlined above. It is probably the phase of work that is most arcane and mysterious for audiences of qualitative reports and requires some account of how conceptual analysis of qualitative data is actually done.

Whatever method of data collection is used in qualitative methodology, it results in raw material in the form of “text” – written words. The first stage of processing text is to code and classify. The investigator therefore reviews his or her textual material for coding in terms of concepts and categories that may emerge from the material or that inform the investigator’s study to begin with. This transformed, coded material can then be regrouped or indexed to facilitate further analysis. This stage of grouping and indexing coded material has been considerably eased by the development of computer packages such as Ethnograph, although such technology cannot replace the far more important human interpretation of text before mechanical manipulation.

The next and most difficult stage is the analysis of both the original textual data and the transformed conceptual material. This phase is the most difficult to prescribe or indeed describe since it entails a large amount of creative interpretation of evidence. Thought processes include constant comparison of evidence regarding different settings or viewpoints represented in the data and the search for deviant or contrasting observations. One very common feature of this central analytical phase is the development of typologies that convey the range of views, responses, or arrangements under study. Indeed a typology may be one of the central results of a qualitative analysis. Thus in the example cited earlier of a study of patients presenting headaches to neurological clinics a key insight into the strengths and weaknesses of such clinics came from developing a typology of the different major concerns that motivated patients to seek medical help for headache. Patients’ views of the benefits of such clinics could be largely understood in terms of these different concerns. A quite contrasting study analysing the role of medical audit advisory groups (MAAGs) identified a typology of three different major models of MAAG.22

Often it is simply that certain unanticipated themes emerge from content analysis. Elsewhere in this issue (p 69) is described an important study of stroke survivors’ views of the benefits of physiotherapy.23 Several patients regarded physiotherapy as a source of faith and hope for the future. The authors argue that such insights would not emerge from the current batteries of available patient satisfaction instruments used in quantitative analysis, which concentrate on more familiar dimensions of process and outcome.

A variety of intimidating terms have been developed to describe the logic of qualitative analysis including “analytic induction” and “grounded theory”.24 In different ways such accounts emphasise one common feature – namely, that qualitative analysis is iterative. The investigator goes back and forth between his developing concepts and ideas and the raw data of texts or, ideally, fresh observations of the field.

**Good practice for qualitative research**

The accusation is sometimes made that qualitative research is an “easy option.” This usually takes the form of an invidious contrast between informally acquired impressions and the systematic rigour of quantitative methodology. In reality the difficulties of obtaining and conveying insights that are convincing to relevant audiences are at least as great for qualitative research. This section outlines some of the general principles of good practice for qualitative methods that have been developed to facilitate analyses that are plausible and relevant to policy (box). These principles are further developed in a growing number of more detailed guides to qualitative methodology.25-27

---

**Principles of good practice for qualitative research**

- Theoretical sampling
- Validation
- Conceptual analysis

---

**SAMPLING**

In quantitative studies considerable attention is given to obtaining a sample which is statistically representative of the population of interest so that generalisations may be made from the study. This is generally achieved through random sampling. In qualitative studies numerical generalisations are less important than *conceptual* generalisations. In this case what it is important in drawing a sample is to ensure that it contains the full range of possible observations so that the concepts and categories developed provide a comprehensive conceptualisation of the subject. The way in which this is done is through “theoretical sampling.”28 On the basis of his or her theoretical understanding the investigator determines what factors might affect variability in the observations and then endeavours to draw the sample in a way which maximises the variability. This may be decided in advance so that, for example, in a study of a clinic the investigator may be careful to interview patients of varying age, sex, social class, and ethnic background or patients seen by different doctors at different times of day. Theoretical sampling may also require adjusting sources of observation in the course of a study to respond to unanticipated patterns or subgroups of experience which may need increased representation in the sample.

Because qualitative data are more cumbersome to manipulate and analyse most qualitative studies are restricted to a small sample size which is unlikely to be statistically representative of the population. Nevertheless, accounts of qualitative research should always provide a clear account of sampling strategies.
to allow readers to judge the generalisability of the conceptual analyses.

VALIDITY
In the same way every effort is required to establish the validity of analyses based on qualitative material. The two most commonly cited methods of validating qualitative analysis have already been mentioned in relation to case studies – that is, triangulation and respondent validation. In the first method every effort is made to obtain evidence from as diverse and independent a range of sources as possible. This approach is not very different from the process of establishing construct validity for quantitative measures in that one is looking for patterns of convergence between data sources that together corroborate an overall interpretation. Respondent validation requires that investigators obtain subjects’ reactions to their analysis and incorporate such reactions into a more complete analysis. However, neither method provides a perfect solution to the problem of validation. Unfortunately, establishing the plausibility of analyses of social settings and organisations such as health care cannot be done by mechanical use of procedures such as triangulation or respondent validation. The world of health care is particularly complex and interpretations of decisions, behaviour, and arrangements will generally vary among participants. Participants may have various reasons for not agreeing with analyses of their behaviour, and, indeed, such disagreements may provide further revealing evidence of how an organisation works.

It is therefore important that users of qualitative methods adopt a number of principles to convince audiences of the validity of their analyses. Have investigators sampled the diverse range of individuals and settings relevant to their question? How much have they drawn on and collected evidence in terms of interviews, records, field notes, that are in principle capable of independent inspection by others? How much have they drawn on, whenever available or appropriate, quantitative evidence to check or test qualitative statements? To what extent do investigators seem to have sought out observations that might contradict or modify their analyses? It remains a matter of judgement for audiences of qualitative studies to determine how systematically investigators have approached such questions in assembling and interpreting their evidence. It is essential that qualitative researchers make more explicit the methods whereby their analyses have emerged so that audiences can make informed judgements about plausibility.

UNDERSTANDING ISSUES OF QUALITY OF CARE
What is the role of qualitative research in understanding issues of quality of care? On the one hand, qualitative methods have enormous potential to illuminate how health care currently operates and the impact of care on patients. Thus puzzling issues such as the persistence of variations in clinical practice should be addressed by qualitative methods. How patients experience and benefit from health services also need examination by such methods. On the other hand, as more systematic evidence accumulates of effective interventions and appropriate forms of health care, another kind of application of qualitative methods is required. The changes in health services needed to promote quality are organisational and cultural and involve differing perspectives of a diverse range of health professionals, managers, and patients. Mechanisms of change such as clinical audit, quality assurance, and the adoption of clinical guidelines are social processes with meanings to participants that we need to understand. In such a complex environment qualitative methodology will be essential to give us models of how organisations change and innovate to adopt quality in health care.

There is an exciting range of qualitative methods capable of providing basic understanding of the processes and outcomes of health care. Rigorous and transparent attention to methodology is needed to convince audiences of the value of insights into the “black box” of health care. An understanding of these methods will in turn promote a fuller appreciation of the insights they can provide.