

Examining patient perceptions of quality care in general practice: comparison of quantitative and qualitative methods

J REES LEWIS

V WILLIAMSON

SUMMARY

Background. *The consultation satisfaction questionnaire and surgery satisfaction questionnaire, and the critical incident technique have been identified as examples of, respectively, quantitative and qualitative (interview) techniques with considerable theoretical merit regarding the measurement of patients' views in a general practice context.*

Aim. *This study set out to assess these techniques in terms of ease of administration and analysis, respondent acceptability, and the extent to which the information provided was useful to the practitioner/practice manager, as well as validity.*

Method. *Patients from three practices completed the interview and questionnaires. Data were provided for each practice giving their own results as well as data from the other two practices and the results of previous research.*

Results. *Both methods were, in the main, received positively by general practitioners, managers and patients. Patient responses to the questionnaires in general followed predictable patterns, variations from which suggested practice-specific problems.*

Conclusion. *There are caveats regarding the use and interpretation of both methods, of which potential users should be aware. This is particularly the case with the consultation satisfaction questionnaire, scores on which, it is suggested, may be on a downward trend over time. It is possible that results from the consultation satisfaction questionnaire/surgery satisfaction questionnaire could be merely demoralizing for practice staff in some instances. Other research supports this notion of demoralization which, although unproven, would reduce the instrument's potential for comparison between studies, and which is, therefore, a finding which requires further attention. Increasing patient expectations are implicated in this.*

Keywords: *quality in general practice; patient satisfaction; assessment techniques; research methodology; comparative studies.*

Introduction

THIS research was carried out at the instigation of three community health councils, and motivated by concerns that those involved in the provision of general practice services are being increasingly encouraged to measure users' views, but are infrequently told how to do so. In essence, the question asked of this

project was: How might a small scale, local, research initiative best measure patients' perceptions of quality care in a general practice context?

The two primary criteria, which guided all stages of the project, were validity — the extent to which information provided by different methods accurately reflected patients' concerns — and utility — the extent to which methods provided information which would be useful to general practitioners and practice managers.

A literature review identified methods which appeared most likely to meet these criteria.¹ These methods produce data of two broad types: quantitative (numerical) and qualitative (narrative). The former are most often survey results gathered in a questionnaire format; the latter, most commonly, by interview. Both have advantages and disadvantages. There are formats though, of both types, where attempts have been made to reduce the disadvantages. Two methods, one quantitative and one qualitative were selected for comparison in this study. More detailed reviews including descriptions of other methods are published elsewhere.^{1,2}

Consultation and surgery satisfaction questionnaires

The consultation satisfaction questionnaire and surgery satisfaction questionnaire were designed to obtain general practice patients' expressed areas of concern, with appropriate regard for reliability and validity.³⁻⁵ The former concentrates on the consultation itself, the latter on the practice in general. They are relatively short, easily understood and quick to complete; each question is clearly focused and, responses express agreement/disagreement on five-point scales which are more discriminating than yes/no responses. Final scores are arrived at by the responses to more than one question, a more robust method than simply using single scores. Response bias and ceiling effects are countered by the inclusion of reversed items (agreement indicates dissatisfaction rather than satisfaction) and overstated items ('I am totally satisfied with my doctor').

The consultation satisfaction questionnaire has four subscales: general satisfaction, depth of relationship, professional care and perceived time. The surgery satisfaction questionnaire has six subscales: general satisfaction, accessibility, availability, medical care, continuity of care and premises. The author of the questionnaires provides data from over 80 surgeries in the Bristol area in which the questionnaires have been used, allowing comparative possibilities for researchers who generate new data with them.

Critical incident technique

Interview surveys provide narrative rather than statistical data, often involve smaller samples or greater investment of time and money than questionnaire surveys, but can yield specific suggestions which have been found valuable in previous research in health and non-health contexts.⁶⁻⁸ The aim of the critical incident technique is to establish what aspects of a clinical encounter determine the patient's impression of it.

The interview covers five stages — pre-entry (appointment): making an appointment; pre-entry (access): getting to, and into, the surgery; entry: the reception and waiting process; appoint-

J R Lewis, BA, research officer and V Williamson, BA, MA, director, Health and Social Policy Research Centre, University of Brighton. Submitted: 14 December 1993; accepted: 23 September 1994.

© *British Journal of General Practice*, 1995, 45, 249-253.

ment: the consultation itself; and exit: from the end of the consultation, including dispensing and follow up.

Patients describe retrospectively their experience of each stage of the process, and are encouraged to say what, in terms of actual observations, incidents, behaviours, or absences of them, caused them to form their opinion. These spontaneous comments are recorded as critical incidents. There are also a limited number of topics which patients are asked about if they do not refer to them spontaneously. These are recorded as prompted incidents and are given less weight in the analysis than critical incidents. Finally, other pertinent comments which indicate expectations, opinions or suggestions, are listed.

The report containing the results of the critical incident technique survey can be presented in a manner which highlights aspects of good practice, and warns about potential irritants without being personally critical of practice staff.⁹ An appendix is included with the report, which lists all the patients' comments, for each stage of the interview.

Method

After consultation with the East Sussex Family Health Services Authority locality managers and the county demographer, a series of practices were invited to participate in the research which took place in March and April 1993. The selection of practices was intended to cover as broad a range of variables as possible and, across the practices included in the study, to be as typical as possible of the population of the area. Of six practices approached, two declined. Of the remaining four, three were selected to participate in the study. These included rural (dispensing) as well as urban practices, from economically mixed areas, housed in diverse premises, fundholding and non-fundholding, single practitioner and larger practices (up to four partners). The research, completed over two- to three-week periods in each practice, included combinations of surgeries to involve the patients of all the partners in each practice.

Consecutive patients aged 16 years or over (or the parents of younger patients) were approached by a researcher in the waiting room after they had seen receptionists, and were given an explanatory letter or a verbal explanation of the research. After their consultation, they were asked if they were prepared either to complete a questionnaire or to give an interview and were assured of confidentiality. Questionnaires were usually completed in the waiting rooms. All interviews were carried out in a separate room. The aim was to interview 30 patients and receive questionnaires from 100 patients in each practice. For reasons of cost effectiveness it was decided to recruit a sample of patients from one practice by mail, to allow for the inclusion of views from a small branch practice, open for two surgeries per week. All patients who attended this branch during the period that researchers were in the main surgery were contacted.

Comparative data, produced by the originators of the consultation satisfaction questionnaire and surgery satisfaction questionnaire are provided to researchers with the questionnaires, and were provided to the practices in their reports. In this study, this was augmented with the results of more recent research produced in practices in the Manchester area. The mean scores on the questionnaires for the three study practices (K, L, E), plus the scores from a fourth practice (W) participating in a different study but collected by the same researchers over the same period (Williamson V, Lewis JR. *The assessment of patient perceptions of quality of care in general practice*. University of Brighton, 1993), were also provided for each practice. The motivation to provide these extra data was simply to make the study as useful as possible to those involved, by giving as much relevant data as possible for comparison.

Reports for each practice comprised a description of the

methods as well as results, including comparisons (*t* tests) across sex and age (with age split at 45 years). Age is the demographic variable which is most consistently associated with variations in expressed satisfaction, older patients tending to indicate that they are more satisfied than younger ones, while sex is less consistently related to differences in expressed satisfaction.¹⁰ Debriefing sessions with general practitioners and practice managers included informal interviews regarding their reactions to the reports.

Results

Consultation and surgery satisfaction questionnaires

The overall mean scores for both questionnaires, together with the range of scores from previous research, are shown in Table 1. Most striking is the fact that two of the practices in the present study scored below the entire range of the comparative data on two of the consultation satisfaction questionnaire subscales.

Age differences. Older patients tended to be more satisfied than younger patients in three of the four practices in this study (Table 1). The results for practice E differ from the common pattern in that there was only one significant difference, and on the consultation satisfaction questionnaire younger patients scored slightly higher than older patients on every subscale. This lack of statistical significance is of clinical significance representing, as it does, an unusual pattern of relationships with patients of different ages. As overall scores for practice E were relatively low, low scores from older patients, rather than high scores from younger patients, were indicated in this practice.

Sex differences. There were no significant differences between the sexes in expressed satisfaction in practices E and W, and only two in practice L: mean score for general satisfaction on consultation satisfaction questionnaire 83.5% for 71 women patients and 81.0% for 29 men patients ($P<0.05$), and mean score for availability on surgery satisfaction questionnaire 39.6% for women patients and 50.3% for men patients ($P<0.05$). In practice K however, men patients expressed greater satisfaction on every subscale and the difference was significant in most cases: on consultation satisfaction questionnaire mean score for general satisfaction 65.5% for 74 women patients versus 76.3% for 26 men patients ($P<0.05$), for depth of relationship 51.4% versus 63.1% ($P<0.01$) and for professional care 67.0% versus 74.2% ($P<0.05$); on surgery satisfaction questionnaire mean score for general satisfaction 53.5% versus 65.7% ($P<0.01$), for availability 43.7% versus 55.0% ($P<0.05$), for medical care 56.1% versus 68.0% ($P<0.01$) and for continuity of care 43.0% versus 54.6% ($P<0.05$). This would suggest that the service provided by practice K was less well liked by women patients.

Critical incident technique

The appointment itself (the most important part of the process) drew most comment within every practice. The second stage — pre-entry (access) — was an issue for few patients. However, if there were problems sufficient to prevent access for some patients, these would not have been identified by sampling among practice attenders only. The predominant themes of the interviews were as follows:

Pre-entry (appointment). Ease/difficulty with which an appointment could be made. How soon it could be arranged. How convenient it was. Ease/difficulty of contacting the surgery by telephone. Ease/difficulty of making an appointment with a particular doctor. The relative merits of open and appointment based surgeries.

Entry. Contact with receptionists. Length of waiting times. Aesthetic and functional aspects of the waiting room. Systems for calling patients in turn.

Table 1. Overall data from the consultation satisfaction questionnaire/surgery satisfaction questionnaire for four practices and by age of respondents, together with data from previous research in Manchester (Kirton C, Foxhall M, personal communication) and Bristol (Baker R, personal communication).

Age of respondents (years)	Mean score for practice (%)				Range of scores from previous research (%)
	K (n = 68/32/100)	L (n = 63/37/100)	E (n = 49/57/106)	W (n = 118/93/211) ^a	
<i>Consultation satisfaction questionnaire</i>					
General satisfaction					
<45	64.1	80.3	68.5	73.1	
45+	77.3**	86.9*	65.9	82.7**	
Overall	68.3	82.8	67.1	77.6	63.5 – 89.3
Depth of relationship					
<45	50.5	66.5	58.4	60.3	
45+	62.8**	76.8*	56.4	74.0**	
Overall	54.5	70.3	57.3	66.4	60.2 – 85.9
Professional care					
<45	66.5	80.2	67.9	73.9	
45+	73.9*	86.0	65.9	82.3**	
Overall	68.9	82.3	66.8	77.7	71.0 – 91.9
Perceived time					
<45	57.0	70.4	69.9	63.1	
45+	61.2	76.0	68.4	70.3*	
Overall	58.3	72.4	69.1	66.3	58.0 – 84.0
<i>Surgery satisfaction questionnaire</i>					
General satisfaction					
<45	54.5	66.7	64.6	71.4	
45+	61.2	73.7	68.6	80.0*	
Overall	56.7	69.3	66.8	75.3	54.0 – 83.0
Accessibility					
<45	71.4	70.8	70.8	72.0	
45+	76.2	64.4	75.6	74.3	
Overall	72.9	68.4	73.4	73.2	59.0 – 82.1
Availability					
<45	43.8	40.4	64.9	59.4	
45+	52.5	46.6	74.8**	66.4*	
Overall	46.7	42.7	70.2	63.8	29.7 – 80.8
Medical care					
<45	57.4	71.2	58.2	64.7	
45+	63.1	74.7	61.7	75.4**	
Overall	59.2	72.5	60.1	69.6	56.5 – 78.0
Continuity of care					
<45	44.0	44.8	74.3	51.4	
45+	50.3	57.0**	74.5	62.7**	
Overall	46.1	49.1	74.4	55.9	32.1 – 82.2
Premises					
<45	38.9	81.7	79.8	76.5	
45+	40.6	77.8	77.4	82.6**	
Overall	39.5	80.3	78.5	79.5	28.2 – 87.4

n = number of respondents: aged <45 years/45+ years/total in each practice. ^aLarger sample size accounts, in part, for the greater significance that can be attached to differences between subsamples in this practice. Patients aged <45 years versus 45+ years: *P<0.05; **P<0.01.

Appointment. Doctor's manner and interpersonal qualities. Doctor's perceived technical competence. Provision of explanations, advice and information. Length of appointment (not absolute length; people do not necessarily want to spend longer than they have to with their doctor, more critical was the feeling of not being rushed, of having enough time).

Exit. The manner in which the consultation ended. Follow up. The availability of a duty doctor for emergencies. Other services within the practice particularly dispensing (where relevant).

Specific successes and problems were also raised in every practice, for example: uncomfortable chairs in one waiting room; consultations in one surgery could be heard in the waiting room;

a doctor, on entering the building, not greeting patients; the importance of a specific invitation to come back if necessary.

Debriefing

Comments (negative and positive) from doctors and practice managers made during the debriefing sessions regarding their reactions to the reports are listed in Appendix 1.

Discussion

This research was conducted on a scale small enough to represent that which could be instigated without too great an investment (in financial or human resources) by a single practice, or by local workers. Although researchers should be perceived as inde-

pendent of the practice, and staff such as receptionists should not be involved in data collection, they can play an important role in facilitating the process by giving an explanatory letter to patients on arrival at the surgery. No one reported that the research disrupted the day to day running of the practices.

Most patients completed both questionnaires within 10–15 minutes. Each interview took, on average, 20–25 minutes with an additional 10 minutes or so for the researcher to complete the notes. Sufficient information to train interviewers, and to cost, run and analyse both methods are reported elsewhere.^{9,10} As with all structured interviews, the critical incident technique requires more training for researchers than the questionnaires but not to a prohibitive extent.

The questionnaires were designed and, in general, appear appropriate for self-completion. The language is simple and the format straightforward; few patients (only those with reading or language difficulties) needed assistance. Despite 'survey fatigue' among some patients who had recently had contact with hospital services and reported completing questionnaires, patients were generally positive about having their views sought. The only frequent criticism made by respondents referred to the repetitive nature of the questionnaires. This is a result of the presence of reversed items, but on very few occasions were potential respondents' objections sufficient to result in refusal to complete the questionnaires. The strengths of this feature, in terms of improved validity, might be judged to outweigh the irritant of repetitiveness. An interview format, being less rigid, is unlikely to suffer similar problems. In the case of the critical incident technique some structure is imposed but the content is dictated largely by the interviewee. Respondent acceptability is therefore unlikely to be a difficulty. Such was the case.

Both questionnaire and interview formats were generally well received by practitioners and managers. Two practices had previously attempted to elicit patient feedback; one with a suggestion box, the other with their own questionnaire. The reports based on both methods employed here were judged substantially more informative. This applies particularly to the critical incident technique, although two practitioners commented that, if anything, the critical incident technique report was too 'gentle' and they felt they were being patted on the back. This was in contrast to their reactions to the consultation satisfaction questionnaire/surgery satisfaction questionnaire.

Just as patients remember a few salient pieces of information given in a consultation,¹¹ so some of the doctors registered, and reacted to, a small number of findings in each report. In the critical incident technique reports, these appear often to be discovered by scanning the appendix, which more than one general practitioner found particularly useful.

The strength of the two questionnaires is that they provide easily digested, simple but salient, findings. General practitioners often focused their attention most on these. They were uniformly judged useful, and interesting.

To the extent that it allows respondents to set the agenda, the critical incident technique should be a valid measure of patients' views, although two of the practices felt that they were doing well on the basis of the critical incident report, but doing badly on the basis of the consultation satisfaction questionnaire/surgery satisfaction questionnaire report. There is a possibility that the critical incident technique can present a slightly rosy picture, but its intention is to highlight and encourage 'good practice', rather than being a survey of complaints.⁹ There is also the potentially confounding factor (reflecting small sample size) of overemphasis on one or two individuals' experiences.

Even so, the results from the interviews can serve as some validation for the questionnaires. Nearly all patient comments on the critical incident technique about the consultation itself could be

grouped into the categories 'interpersonal skills', 'technical competence', 'giving of advice/information', and 'length of consultation'. These tally closely with the subscales on the consultation satisfaction questionnaire: 'depth of relationship', 'professional care' which include questions relating to perceived technical competence and to information giving, and 'perceived time'. For the practitioner, this may constitute a more clinically significant validation for the questionnaires than statistical factor analysis (Kinnersley P, personal communication).

One of the principal advantages of a standardized instrument is the possibility of comparison with data produced in other research. This is dependent, in part, on the instrument's reliability, but also on the stability of the population from which samples are drawn, and in this respect there are, perhaps, issues to address for potential users. The information provided with the questionnaires includes mean scores from over 80 practices in which their author has used them. There is obviously a time lag (18 months to two years in this case) between the production of different sets of data, and the impression gained from researchers, not only in this project but also in Cardiff (Kinnersley P, personal communication) and Manchester (Kirton C, Foxhall M, personal communication) is that the range of scores being produced on the consultation satisfaction questionnaire is getting lower. In the present study mean scores for two practices fell below the entire range of previous scores on two out of four subscales. This would have been three subscales but for the inclusion of the Manchester data. It should be said that these practices did not give researchers the impression of providing a service sufficiently poor to generate exceptionally low scores. Nor did the critical incident technique interviews indicate problems as severe as these scores imply.

It is possible that regional, or sample, differences could account for this impression of declining scores. This seems unlikely though as a variety of practices was included in each piece of research; and there is no visible trend which would support such an interpretation. In the present study the two low scoring practices were dissimilar in many respects. Practice K is in an urban area, mixed economically, with several doctors working out of an old building. Practice E is a single-handed practice in a prosperous rural area, offering many other services (a woman doctor available for one surgery per week, counselling, acupuncture, and so on), from a purpose built surgery.

One explanation is that satisfaction scores have decreased, or are decreasing, over time. As expectations and satisfaction are clearly linked,¹² this leads to the suspicion that, in respect of the consultation itself, patients' expectations are rising. A recent report from the Association of Community Health Councils of England and Wales (1993) opines that, since the introduction of the patient's charter, 'patient expectations are on the increase and this is having a substantial effect on many workloads'. A similar opinion was voiced by members of three of the practices in the present research.

There is a potential, therefore, if comparisons are not time specific, for questionnaire data provided without corresponding qualitative analysis, to exacerbate the feedback loop comprising rising patient expectations and decreasing general practitioner morale. A situation such as this clearly weakens the potential to make useful comparison between different studies using these, or any quantitative, measures of patient satisfaction.

The impression that scores are decreasing over time is a hypothesis suggested by the data rather than a conclusion. However, these findings may well reflect changing circumstances (higher patient expectations) rather than inadequacies in the design of the instruments (or in the practices). This is not so much criticism of the instruments themselves as a caveat against their inappropriate use, but worthy of note as these instruments have been widely distributed.

Appendix 1. Comments from general practitioners and practice managers.**Critical incident technique**

'Gives concrete examples regarding things like having a female doctor, or open surgeries. We thought people liked open surgeries but weren't sure how much... This will give us confidence to ensure that these services remain.'

'Have already made small changes' (for example, greeting patients on entering surgery building).

'Sometimes a particular patient is a problem and they are never satisfied... This patient may bias the comments in what is a relatively small percentage of the practice.'

'It was useful to know what makes the patient feel welcome and what the patient appreciates about a small practice. We can therefore try and concentrate on these factors and improve them where necessary.'

'In particular the section about the consultation and the doctors' behaviour was very helpful. It is always hard for the doctor to know how patients perceive him. They are not likely to tell him and there is not usually anyone else present during the consultation. Even if the staff get feedback they are in no position to give him a straight answer.'

'We also gained some useful tips on what the patients appreciate and how they perceive the service to work efficiently, for example, contact from surgery to follow up their treatment, being informed promptly about test, whether clear or not.'

'Helpful, didn't necessarily identify new problems, but confirms suspicions and gives the practice manager "clout".'

'Useful, for example to argue with the FHSA.'

'Interviews: much more effective, particularly the appendix listing comments.'

'Felt a lot better after reading this report [as compared with the consultation satisfaction questionnaire/surgery satisfaction questionnaire].'

Consultation and surgery satisfaction questionnaires

'Bland categories — uninformative.'

'Makes you think about different aspects of care.'

'...there are no clues as to what leads a patient to his or her opinion... If the patients are not satisfied the questionnaire analysis did little towards suggesting why.'

'It was helpful to compare our survey with others, especially the other three practices which were questioned at the same time by the same researchers. This seemed much more relevant than comparison with studies done elsewhere and at a different time. Unfortunately this comparison would not be available if studies were done on a one off basis.'

'Although comparisons are helpful, they are likely to lead to considerable stress when a practice finds it is low on the list of patient satisfaction and there are no clear reasons as to why.'

'Can be useful, not demoralizing.'

'Revealed things one is intrinsically aware of — but not conscious of.'

'Highlights problems.'

'Would be better used by an FHSA than by a single practice.'

'...is useful, even to have what you think confirmed.'

'Very interesting, more interesting [than the critical incident technique].'

References

- Lewis JR. Patient satisfaction with general practice services: literature review. *Soc Sci Med* 1994; **39**: 655-670.
- McIver S. *Obtaining the views of users of health services*. London: Kings Fund Centre, 1991.
- Baker R. Development of a questionnaire to assess patients' satisfaction with consultations in general practice. *Br J Gen Pract* 1990; **40**: 487-490.
- Baker R. The reliability and criterion validity of a measure of patients' satisfaction with their general practice. *Fam Pract* 1991; **8**: 171-177.
- Baker R, Whitfield M. Measuring patient satisfaction: a test of construct validity. *Qual in Health Care* 1992; **1**: 104-109.
- Pryce-Jones M. Not how many but why? A qualitative approach to customer relations. *Health Service Management* 1988; **84**: 175-177.
- Gau D, Pryce-Jones M, Tippins D. Satisfactory practice. *Health Service J* 1989; **99**: 1464-1465.
- Bradley CP. Turning anecdotes into data — the critical incident technique. *Fam Pract* 1992; **9**: 98-103.
- Callaghan D, Caple T. *Managing customer relations*. London: North West Thames Regional Health Authority, 1986.
- Hall JA, Dornan MC. Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. *Soc Sci Med* 1990; **30**: 811-818.
- Ley P. Improving patients' understanding, recall, satisfaction and compliance. In: Broome AK (ed). *Health psychology: processes and applications*. London: Chapman and Hall, 1989.

- Linder-Pelz S. Social psychological determinants of patient satisfaction: a test of five hypotheses. *Soc Sci Med* 1982; **16**: 583-589.

Acknowledgements

Data from, and the informed comments of C Kirton and M Foxhall, Manchester MAAAG and of P Kinnersley, Department of General Practice, University of Wales College of Medicine, Cardiff, are gratefully acknowledged. Perceptive and encouraging comments from R Baker are also gratefully acknowledged. We are grateful for the continued support and advice from steering group members representing Brighton, Eastbourne and Hastings CHCs, East Sussex FHSA, and local PPG and GP representatives. We thank those who collected data, and staff and patients at all participating practices.

Address for correspondence

Ms V Williamson, HSPRC, Department of Community Studies, University of Brighton, Falmer, Brighton BN1 9PH.

RCGPResearch
Network

The RCGP is pleased to offer two appointments as RCGP Research General Practices starting on 1st October 1995 for up to two years.

Applications are invited from members of the RCGP who are working within a general practice in the United Kingdom which has a strong interest in, and an established track record of undertaking successful research in general practice and whose practice wishes to maintain its research work in general practice.

The purpose of the College's award is to recognise the additional costs incurred by practices in order to provide the infrastructure necessary to enable research to be undertaken within general practice. Funding of £12,500 per year will be available for two years, subject to satisfactory review.

Further details and an application form can be obtained from Professor Denis Pereira Gray, Chairman of Research, Royal College of General Practitioners, 14 Princes Gate, London SW7 1PU, to whom applications should be submitted by 1st July 1995.