

Acceptability of methods and measures used to determine quality of general practice consultations: results of a focus group study and an acceptability questionnaire

Margaret Maxwell, David J Heaney, John GR Howie, Jeremy J Walker Department of Community Health Sciences – General Practice, University of Edinburgh, Edinburgh and **George K Freeman** Imperial College School of Medicine, Department of Primary Health Care and General Practice, Chelsea and Westminster Hospital, London, UK

General practitioner activity is increasingly under pressure to monitor its performance. The involvement of service users in the development and assessment of services is said to be a key feature of this process. This article reports on the acceptability among general practitioners of a patient-completed post-consultation measure of outcome (the Patient Enablement Instrument; PEI), and its use in conjunction with two further indicators of quality, namely time spent in consultation and patients reporting knowing the doctor well. The survey was conducted using focus groups and the administration of a postal questionnaire among a group of general practices that had participated and received feedback from a large quantitative study testing these measures. The focus group study provided useful insights into general practitioners' perceptions of patient assessment of their performance and their concerns surrounding the measurement of general practice activity. The general practitioners' perceptions of the measures under the study were enmeshed within these concerns overall. The PEI was seen as being generally acceptable as a measure of patient assessment of care, and the methods of data collection were acceptable for routine use in general practice. General practitioners who performed better in terms of their feedback scores generally approved more of the proposed measures. However, these general practitioners were not comfortable with the concept of assessment of the clinical interaction by patients, and were anxious to link such assessment explicitly with clinical (disease-related) outcome. Doctors who performed 'better' were no more likely than those who performed less well to advocate more use of patient assessment, or to believe that patient assessment of consultations is a reliable quality indicator. These concerns need to be addressed if patients' assessments of their care are to be taken seriously.

Key words: focus groups; general practice; general practice assessment; patient-centred measures of outcome; primary care; quality of care

Introduction

The introduction of clinical governance and the move towards continuing assessment of general

practitioners' performance have placed the need to find outcome measures in general practice in a new light. There are currently several approaches to the measurement of achievement in terms of improv-

Address for correspondence: Margaret Maxwell, Department of Community Health Sciences General Practice, Primary Care Research Group, University of Edinburgh, 20

West Richmond Street, Edinburgh EH8 9DX, UK. Email: Margaret.Maxwell@ed.ac.uk

©Arnold 2002

10.1191/1463423602pc075oa

ing quality and effectiveness in general practitioner services. One approach is that of using targets and performance indicators which are linked to payments within the general practitioner's contract and are mostly practice-based incentives (Campbell *et al.*, 1998; Elkan and Robinson, 1998). Another approach involves developing specific outcome tools which measure the quality of general practitioner care. The latter approach can take two forms, namely the development of clinical standards as measures of quality (focusing on technical competence) (McKee, 1997; Cantrill *et al.*, 1998; Campbell *et al.*, 1999; Baker, 2000; McColl *et al.*, 2000), or the development of patient-centred measures of quality. Patient-centred measures of quality can range from satisfaction outcome measures (Locker and Hunt, 1978; Wolf *et al.*, 1978; Cleary and McNeil, 1988; Baker, 1990; Fitzpatrick, 1991; Williams, 1994; Kinnersley *et al.*, 1996; Cleary, 1998; Steven *et al.*, 1999) through to patient assessment of a range of aspects of their care, such as accessibility and availability, technical care, and interpersonal care (Baker, 1991; Flocke, 1997; Gelb Safran *et al.*, 1998; Roland *et al.*, 1998).

Several measures are currently being tested and validated for use in general practice and primary care in the UK. One such instrument, namely the Patient Enablement Instrument (PEI) (see Box 1) was developed as a patient-completed post-consultation measure of outcome based on the concept of 'enablement' (Howie *et al.*, 1997; 1999). In developing this measure the emphasis was on deriving an instrument which would reflect patients' perceptions of the effectiveness of their consultation, with particular reference to increased feelings of understanding about their health and illnesses, and to their feelings of ability to cope. This instrument has been further tested in a large quantitative study in four areas of England and Scotland, namely West London (Ealing, Hammersmith and Hounslow), Coventry, Oxfordshire and Lothian (Edinburgh). This testing was carried out in 53 practices and included 221 doctors. General practitioner and practice details were collated prior to conducting the main survey, which was a cross-sectional study of all consultations ($n = 30\,582$) during a 2-week period in March/April 1998. Data collected at consultations included a patient-completed preconsultation questionnaire which asked about their reasons for consulting, whether they would see their usual doctor and how well

they thought they knew the doctor they were seeing, and some expectations of the consultation, as well as a measure of social and psychological well-being. Doctors recorded the time spent in consultation and then patients completed a brief exit questionnaire, which included the PEI. Full details of this study have been published elsewhere (Howie *et al.*, 1999).

As part of this larger study testing the PEI (as an outcome measure) and other quality indicators (time spent at consultations and continuity as measured by how well patients reported knowing their doctor) among different general practice populations (Howie *et al.*, 1999), we also set out to determine the acceptability of this work to participating general practitioners and practice staff. The survey covered the acceptability of the data collection methods in general, as well as asking about the credibility of the individual measures being tested and the future potential for these measures to be used as part of routine assessment of the quality of care at general practice consultations. The survey was conducted using focus groups and a postal questionnaire, and this paper presents the results.

Methods

The focus groups

Participants had received feedback results twice during the course of the main study, and an appropriate methodology was sought that would capture the impact of these results whilst also broadening the focus of discussion beyond personal scores. We aimed to establish debate among the participants with regard to our methods and the potential uses to which they could be put. Levy (1979) reports the advantages in hearing how focus group participants respond to each other, and how this gives insight not just into their natural vocabulary on a topic, but also into when they are willing to challenge others and how they respond to such challenges. The use of focus groups within general practice research and their advantages have also been documented elsewhere (Barbour, 1995; Kitzinger, 1995).

All 221 general practitioners who had taken part in the quantitative survey and their practice reception staff were invited to a local feedback/discussion meeting. There was sufficient homogen-

city in subjects' backgrounds to exclude the necessity to run separate groups for different categories of participants (Kreuger, 1994). However, where possible we separated partners from within the same practice and other practice staff from their GP employers. Techniques to minimize the level of moderator involvement (Morgan, 1988; 1993) were employed in order to avoid questions about the main study and its results being addressed to the moderators.

Seven focus groups were conducted in four regions (Oxford = 2, Coventry = 1, London = 1 and Edinburgh = 3), involving a total of 56 participants. This represented 41 practices covering a range of sizes and levels of deprivation. Group sizes ranged from 7 to 10 participants. Participants included 46 general practitioners, 9 practice managers and 1 receptionist. Each focus group had a moderator and an observer, and the transcriber also spent some time in each group in order to familiarize themselves with the speakers so as to facilitate accurate transcription of the tape-recorded sessions. A copy of the topic guide containing four questions was given to each participant (see Box 2). Each session lasted for approximately 45 minutes.

Analysis of focus group discussions

All focus group discussions were fully transcribed. Transcripts were analysed to establish the main categories (themes), and to relate these to the original research questions. These main categories were further explored in order to determine subcategories and establish conflicting or contradictory views within these categories. Specific note was made of incidences where an expressed opinion sparked consensus, disagreement or a change of opinion among respondents. The majority of this paper focuses on the contributions of the general practitioners. The practice managers' and receptionists' contributions were mainly limited to comments on the methods of data collection, and are reported only under this section of the paper.

The acceptability questionnaire

The most prominent focus group themes and the issues raised within them were used to formulate an acceptability questionnaire for use with all participants in the main study (including practice managers who received separate questionnaires). Only the results of GP questionnaires will be

reported in this paper. Nonresponders to the initial postal questionnaire received a reminder 6 weeks later. Overall, 113 of the 221 GPs who took part in the main quantitative study returned completed 'acceptability' questionnaires (51%). Of the 46 GPs who attended the focus groups, 32 GPs (70%) returned completed questionnaires. Data for the focus group attendees were initially analysed separately from data for those who participated in the main study but did not attend the focus groups. There were no significant differences in responses between the two groups. Therefore the results were amalgamated and are presented for the total GP population.

Before presenting the results, a few words must be said about the background to conducting these focus groups. Several weeks prior to conducting the groups, the GP participants had been presented with the results from the quantitative study. This provided them with a 'score' of their performance on the PEI (as outcome) and two further (process) measures of quality, namely their average length of consultation (time), and whether their patients reported 'knowing the doctor well' (our proxy measure for continuity of care) (Howie *et al.*, 2000). Not all doctors performed well on our measures, and there was the potential for their views on the methods, results and potential applications of the results to be negative. The final section under 'Results' presents correlation coefficients between doctor-level performance on our 'quality' measures and their corresponding answers to the acceptability questionnaire.

Results

We attempted to focus the discussion on the specific tools and methods of data collection, but this led to more general discussion of issues such as GPs' perceptions regarding patient assessment of their performance, general problems surrounding the measurement of general practice activity, and opinions about the incentive/target payment schemes currently in use. General practitioners' perceptions of the PEI and the other two measures of quality are enmeshed within these concerns overall.

There were no apparent regional differences in either the prominent themes which emerged from

the groups or the consensus of opinion surrounding these themes.

General practitioners' perceptions of patients' assessment of quality of care

General practitioners acknowledged the importance and, at most times, the relevance of patients' assessments of outcome. However, there were also several concerns which, once raised, gained support from the majority of general practitioners.

The first of these is the general practitioners' concerns about the ability of patients to determine 'good practice' in medical care. There was general agreement that patients lacked the ability to assess clinical aspects of their care, and concern that an instrument which measured patients' perspectives of their health care (particularly interpersonal aspects of care) would reflect a one-sided viewpoint that did not take into account the clinical competence of the doctor.

No, it didn't mention anything clinical, and I said that to [] right at the beginning, that it's all very well that a patient feels wonderfully enabled, but if you fail to notice that they have X . . .

(F5P2)

Well, it's a feel-good factor, isn't it, it's not a measure of 'they feel good but had a heart attack as they walked out the door'.

(F4P7)

This led to further agreement that patients' expectations and demands could conflict with what would be appropriate or best clinical practice. The use of language across the groups was similar, with many references to 'the consumerist society operating within the health care arena'. General practitioners thought that patients' expectations may be too great or unrealistic, and the example of the patient demanding a prescription was given as evidence of 'bad consumerism and expectation and the demands' of patients (F1P4).

I mean the example of the antibiotics, expecting antibiotics and not getting it, they didn't feel enabled but a few days later they felt better. That doctor was quite right you know, that was just an example.

(F5P2)

I thought, in the consumerist society we live in, I thought, to leave the judgement as such wholly to the perception of a patient, probably is not absolutely right and correct. . . .By that I mean, any consumer you know, no sorry customer rather, not consumer, using any provider, whether it is a supermarket or whatever, they have expectations and they give the marks and if you ask them whether their demands or expectations are met or not, if they're not met then they won't be satisfied.

(F1P2)

However, even within these strongly held viewpoints there was an acceptance that patients could still judge whether their general needs had been met.

The study doesn't include clinical outcomes so the patients could go out very satisfied and drop dead outside the door and you would have failed. But on the other hand, most patients come because they want various things and I think this measure is as good a measure as I've seen. And whether they get what they want, whether what they want is appropriate, whether you know, whether there are underlying clinical problems that are not addressed, I think for most they come with a pretty good idea of what they want and they go away with a pretty good idea of whether they got it.

(F3P4)

In response to patients' lack of clinical insight, some GPs did agree that more time spent explaining GPs' decisions might lead to greater enablement, and so they acknowledged the value in enabling patients.

'Cos in theory, if they are then able to cope in future and don't require a consultation or they feel confident enough, 'cos some of them do come to see us because they don't feel confident in coping, and if you can make them feel confident in coping that is presumably good medicine and good for them. So not meaning to be negative about everything, I think there's very positive prospects in this, it's just the fine tuning.

(F7P4)

This general reluctance to accept patient assessment of the care they received was confirmed in the questionnaire results. Table 1 shows the percentage of doctors who expressed a desire to see more use of patient assessment of consultations, and whether they saw patient assessment of consultations as a 'reliable' quality indicator. Although 67% agreed that more use of patient assessment of consultations should be used to assess quality of care generally, one-quarter (25%) were still unsure about this approach. Only 37% of doctors thought that patient assessment of consultations was a 'reliable' quality indicator; 19% thought that patients' assessment was 'not reliable' and 44% were 'not sure'.

A further area of concern for general practitioners was patients' expectations of the role of the general practitioner and the impact that patient assessment of care might have on the doctor-patient relationship. A significant number of doctors in all groups discussed whether patients' expectations of their role should go beyond the clinical role. They questioned how far GPs should go in meeting patients' expectations, demands and needs, particularly if they were nonclinical needs. They also discussed whether it should be expected that general practitioners have the ability or responsibility to help patients to achieve some of the elements implied within enablement (e.g., 'cope with life').

And the thing which goes along with that, which rated or scored very low is the social

Table 1 Acceptability questionnaire – percentage of GP responders who expressed a desire to see more use of patient assessment of consultations and whether they see patient assessment of consultations as a reliable quality indicator

Item	Response	<i>n</i>	Percentage of total
Do you think more use of patient assessment of consultations should be used to assess quality of care generally?	Yes	73	67.0
	No	9	8.3
	Not sure	27	24.8
	Total	109	100.0
Is patient assessment of consultations a reliable quality indicator?	Yes	40	37.4
	No	20	18.7
	Not sure	47	43.9
	Total	107	100.0

problems, because social problems need longer time. But whether really, going back to what [another respondent] was saying, whether really as GPs we are well placed to deal with social problems and should we deal with them, or do we give this to somebody else like a social worker, like a counsellor or somebody else . . . in an ideal world maybe we should spend more time with patients, but the reality is not like that.

(F1P3)

A: Well, I didn't like the first question on the enablement measure, and I don't feel it reflects how the doctor is enabling the patient and I think it's something out of the doctor's control. And certainly coming from a practice that has quite a bit of deprivation, I feel that was certainly against us.

B: So what was the first question?

A: It was 'I was able to live my life better' or something.

C: Cope with life?

B: Yeah.

C: Yeah, I think that brings in a lot of the patient's own background.

B: Sure it does, carry on.

D: Why is that a problem, why do you feel that's a problem, because it's in an area of deprivation?

A: I feel that we have very limited resources to deal with how patients can cope with their life, they all have social problems that we cannot do anything to help.

(F7P1)

Finally, general practitioners made assumptions about the impact that patient assessment of care (or the achievement of good scores) might have on the doctor-patient relationship. For example, with regard to 'GPs having to be pleasant to patients all the time,' the result of enhancing outcomes for patients would result in doctor-dependent patients.

I think if you did it as a reward scheme all doctors would [have to] be very pleasant to all the patients all the time (laughs).

(F3P13)

And there are some groups that you want to disenable, you don't want to make them doc-

tor dependent, so you want them to go out the consulting room feeling a bit, almost annoyed, saying ‘I went to that doctor and he didn’t do anything for me today’.

(F6P2)

There were some counterarguments to these statements, with the acknowledgement that more time spent explaining things in more detail to patients might help to prevent unnecessary future consultations.

Acceptability of data collection methods

Both GPs and practice managers reported high levels of acceptability of the data collection methods. General practitioners perceived little extra workload for themselves, and very little interference with the running of their surgeries.

The results of the acceptability questionnaire confirmed the opinions expressed within the focus groups. Table 2 summarizes the responses with regard to data collection methods. In total, 96% agreed that methods of data collection were acceptable or very acceptable in terms of the workload for GPs, and 93% said that methods of data collection were acceptable in terms of the workload for practice reception staff. A total of 58% reported that the data collection methods did not affect the running of the surgery, and of the 36% who said

that it did, many qualified this by stating that this was not to any great extent, or that it slightly increased the time of the surgery as they waited for patients to complete forms.

The majority of GPs (74%) were under the impression that most patients had completed the questionnaires. The main problems with completion related to patients with learning difficulties, elderly patients, the poorly sighted who had not brought their glasses, those with difficulty in understanding the English language, and to a lesser degree mothers with young children. There were very few instances of unforeseen problems with regard to data collection, with only 10% reporting problems, the commonest being practice staff forgetting to hand out forms, or when surgeries were running to time patients often had insufficient time to complete the preconsultation questionnaires.

Perceptions of the PEI instrument and its individual items

Face validity

The consensus was that the instrument was generally a ‘good’ measure, ‘a reasonable set of questions’, which had relevance to patients and their care. The measure (or the achievement of the elements contained within it) was also described

Table 2 Acceptability questionnaire – GP responses to questions on data collection methods used in the study

Measure	Percentage of total respondents who replied				n
	Very acceptable	Acceptable	Unacceptable	Very unacceptable	
What was your overall impression of the methods of data collection used in this study in terms of:					
Workload for GPs?	19.6	75.9	3.6	0.9	112
Workload for reception staff?	15.5	77.3	7.3	0.0	110
Measure	Percentage of total respondents who replied			n	
	Yes	No	Not sure		
Did you feel that the data collection methods involved for GPs affected the running of the surgery?	35.7	58.0	6.3	112	
Did you get the impression that the majority of patients completed the questionnaires?	74.3	5.3	20.4	113	
Were there any unforeseen problems with regard to data collection in your practice?	10.4	82.1	7.5	106	

as being ‘worthwhile’ and ‘very worthy’, in that helping people to live with their illness was really a more accurate reflection of what GPs could do for patients compared with curing them.

We are trying to get a flavour of what patients feel about doctors and the consultation, and that’s as far as it can go. If you want more than that you’ve got to ask more specific questions and then the whole thing gets complicated and bogged down. So, I think the whole thing is giving a general impression of how patients relate to their doctors and to the consultation in particular, and that’s probably as far as it can go.

(F2P5)

But it does seem a worthwhile measure, I mean helping people to live . . . as a medical student or someone breaking into the reality of life as a doctor . . . that helping people to live with their illness was much more of what you were going to do than curing people. So it was nice to see those questions . . . quite a helpful measure of what you’re doing. But I mean, it must be a very important part of what we do. I mean if you were scoring well in helping people to cope with life, cope with life’s a bit vague, but cope with your illness, I mean I think so much of it is about keeping people going.

(F3P11)

There was also consensus that the PEI was an improvement on satisfaction questionnaires in general in that it was more relevant to aspects of care that patients would perceive as important. Table 3 shows the quantitative responses of GPs to the PEI. Again they reflect a general acceptance that the questions contained within the PEI reflect outcomes of care which are important to patients (81%), are better at assessing ‘quality’ of care than satisfaction questionnaires (58%), and relate in some way to patients’ assessment of a GP’s patient-centredness (64%).

Interpretation and scope

The focus group participants commented that we currently have little understanding of how different patients will interpret some of the questions (e.g., ‘cope with life’, ‘confident about health’), and whether these questions are appropriate to all

groups of patients. The GPs were specifically concerned with the meanings for those with chronic illness or those with different linguistic and cultural backgrounds to English-speaking patients.

The GPs working among large populations of ‘other-language’ speakers offered some reasons for these potential differences in interpretation and understanding. For example, the question on ‘ability to cope with life’ might appear rather philosophical to some South Asian populations, rather than relating to coping with social, psychological and physical problems, as is implied for English-speaking populations.

Time frame

In relation to concerns about real (health) gain being achieved through enabling consultations, the respondents in all groups referred to the ‘snapshot’ view that this outcome measure offered. It captures patients’ immediate reactions to the consultation, in effect asking them about ‘coping’ before they have had a chance to reflect on this. In addition, they were concerned that the measurement of outcome at a single consultation may not do justice to the long-term relationship and ongoing care of patients within general practice.

but I mean looking at it in a more positive way, I think it’s a very useful attempt to get an instrument that measures patients’ perspectives on health care and the relationship between the doctor and the patient. My big concern with the instrument is, it’s such a snapshot that it doesn’t reflect the complexity of the doctor–patient relationship and probably isn’t a good reflection of the ongoing relationship unless you can validate it.

(F2P2)

However, there was also the pragmatic response that we are measuring *something* at the consultation.

Sure there are other ways of measuring them, but these tend to be outwith the consultation, don’t they? And this gets to the nubbins of people coming to see us and how they go out, how they are dealt with.

(F6P3)

Perceptions of the other measures of quality

While most respondents generally agreed and accepted that more time spent with patients is a

Table 3 Acceptability questionnaire – general practitioners' perceptions of the PEI in relation to whether they reflect outcomes of care that are important to patients, are better at assessing quality than satisfaction scales and whether they relate to patient assessment of GPs' patient-centredness

Measure	Percentage of total respondents who replied				n
	Yes, absolutely	Yes, somewhat	No, not entirely	No, not at all	
Do you think that the PEI reflects outcomes of care which are important to patients?	22.0	58.7	17.4	1.8	109

Measure	Percentage of total respondents who replied			n
	Yes	No	Not sure	
Do you think that the PEI is a better tool for assessing quality than satisfaction questionnaires?	57.9	8.4	33.6	107
Do you think that the PEI could be a measure of patients' assessment of GPs' patient-centredness?	64.2	11.3	24.5	106

good thing and therefore desirable, they defended their use of shorter consultations and questioned whether more time is always used effectively and whether it increases clinical accuracy. They recognized that at the practice level some GPs are more skilled in the technical aspects of clinical care and some have better communication skills (for dealing with psychosocial problems which may require more time), and patients may learn to distinguish between GPs and choose one according to their needs. There were mixed perceptions as to whether GPs really had the ability to alter consultation times and offer longer consultations without the input of resources to begin with.

From the questionnaire data, 76% of doctors generally accepted that average consultation lengths above 9 minutes are better for patients than those below 7 minutes.

'Knowing the doctor well' was also accepted as being a good thing in its own right, and doctors welcomed the 'evidence' which they perceived this study as providing for working with personal lists in smaller practices, as opposed to the policy trend for working in increasingly large units, which reduces the likelihood of knowing one's patients.

Perceptions on use as an inventive/reward scheme

Strategies for providing incentives/rewards to help poorly performing practices to improve were

Primary Health Care Research and Development 2002; 3: 29–41

discussed, and it was proposed that not all schemes should be based on rewarding those who do well.

maybe they could reinvent a new system to actually put more resources in selectively to allow practices that are scoring low both from the patients' perception and clinical indicators point of view, to try and boost the number of staff available to look after patients.

(F7P17)

It was interesting to note that there was little discussion of the possible reasons behind poorly performing practices beyond practice population and patient characteristics (e.g., the suggestion that consultation rates are higher in deprived areas, which affects the GPs' ability to increase consultation length).

In comparison with current payment schemes, some saw our measures of quality as an improvement on current schemes (which they reported as including items which are easier to measure, rather than items which reflect 'quality', and they cited cervical screening and immunization targets as examples). They agreed that they would rather be tested on meaningful measures such as those we proposed than on current target schemes. Others were of the opinion that it was as good a basis for an incentive scheme as anything, but it remained to be seen whether it was better.

- A: There is another area which I just want to toy with, which is that we have currently accepted some performance indicators as being 'performance' indicators, which they are not. Like target payments, it has got nothing to do with GP performance. I would rather be tested on something like this than on my target payments.
- B: Largely what this discussion is about, I feel that the answer to question 4 is, it is as good a basis for an incentive scheme as anything else, and it remains to be seen whether it is better.
- A: I think we'll vote for that, but I think it's better, it's better.
- C: More an area of where you can't get your smears or you can't get your targets, you could still be doing a good job.
- D: Must be better than targets.

(F2P14)

- A: I mean, I think with the provisos that we've been discussing, I think it is a, I think in many ways it could be very useful in incentive-based reward schemes, because unlike so many of the others, it connects much more with what we are actually doing.
- B: Absolutely.
- A: As GPs, seeing patients. And it's also not scuppered, perhaps quite so easily by ridiculous things which are completely out of our control. Like, well you know, the absurd business of the generic prescribing which is absolutely farcical really . . . Or things like, getting cervical smear targets when the fact that somebody just refuses to have one is not allowed for and you may lose your target.

Several respondents: Yeah.

(F7P16)

Doctors believed that incentives/rewards should be based on aspects which were under the control of the GP, and it needed to be made clear what GPs could do to improve. They argued that some GPs might spend time and know patients well but still not be enabling, and we needed to be able to offer education and advice as to how they could improve their skills.

Doctors also agreed that even if our measures

were not linked to a reward or incentive scheme, they still had potential for assessing the quality of care being provided by GPs, and they could be a good means of identifying 'sick doctors' or poorly performing GPs (this had indeed occurred within this study).

Comparison of doctor-level performance on our quality measures with their corresponding results in the acceptability questionnaire

Of the 171 doctors for whom we had complete scores on all three measures of quality, 96 doctors (56%) responded to the acceptability questionnaire. On comparing the results of the process and outcome measures of responders with nonresponders, it was found that responders were better known by their patients than nonresponders ($P = 0.001$), but there was no significant difference in PEI scores or mean consultation lengths between the two groups. In general, we conclude that those who responded to our final questionnaire were not materially different in terms of their performance on our key measures to those who failed to respond.

Many of the items in the acceptability questionnaire had four response options, ranging from very acceptable through to very unacceptable. Responses were scored as 3, 2, 1, 0 (i.e., very acceptable = 3 and very unacceptable = 0). Other questions only had three response options, which have been scored as 'yes' = 2, 'not sure' = 1 and 'no' = 0. Although rank correlations with such data are not ideal, we confirmed the results via Kruskal-Wallis tests, and the findings of the correlation analysis were supported. Table 4 shows the Spearman correlation coefficients between our three key measures of doctor-level performance and several items in the acceptability questionnaire.

Doctors who performed 'better' in the main quality study in terms of higher PEI scores and longer consultation times had more positive perceptions of the acceptability of the data collection methods. Doctors who performed better in terms of PEI scores found the process less disruptive ($P < 0.01$) and had more optimistic impressions of completion rates for our questionnaires ($P = 0.02$).

Those doctors who performed 'better' in the main quality study in terms of higher PEI scores and longer consultation times also had more positive perceptions of the appropriateness of enablement as an outcome measure, and its ability

Table 4 Spearman correlation coefficients between the three key measures of doctor level performance and items contained within the acceptability questionnaire

	Mean PEI score	Mean consultation length	Mean 'knowing doctor' score	<i>n</i>
Acceptability of workload to GPs	0.40 ($P < 0.01$)	0.26 ($P = 0.01$)	NS ^a	95
Data collection affects running of surgery	-0.44 ($P < 0.01$)	NS	-0.20 ($P = 0.06$)	95
Impression of majority of patients completing questionnaires	0.24 ($P = 0.02$)	NS	0.17 ($P = 0.09$)	96
Data collection resulting in unforeseen problems	NS	NS	NS	89
Generally agree that enablement is an appropriate outcome of consultation	0.39 ($P < 0.01$)	0.24 ($P = 0.02$)	NS	93
Generally agree that enablement reflects outcome which is important to patients	0.32 ($P < 0.01$)	0.40 ($P < 0.01$)	NS	93
Willingness to work with us again	NS	NS	0.17 ($P = 0.10$)	95

^aNS, not significant ($P > 0.10$).

to reflect outcomes which are of importance to patients. They were also more likely to view their scores as making sense to them, given their views of their own personal assessment of their working practices.

One of the more interesting findings, which confirms the qualitative findings described above, is that 'better' performing doctors were no more likely than those who performed less well to think that there should be more use of patient assessment of consultations to assess quality of care generally. Neither were 'better' performing doctors more likely to believe that patient assessment of consultations is a reliable quality indicator (all P -values > 0.1). The general reluctance to accept patient assessment of performance still holds even for those doctors who performed well on the basis of patient assessment.

Another way of assessing the acceptability of our methods is to ask if respondents would be willing to participate in similar work again. We did ask this question, and of the 112 doctors who responded to this item, only seven rejected the possibility outright. Again there was no association between doctors' performance on our measures and their willingness to work with us again. We take this as testimony to the fact that our methods are highly acceptable and could be employed for routine use within general practice with minimal disruption to practices.

Discussion

Few studies have gone back to participants to seek criticism of the instruments and data collection methods used, particularly after providing feedback on their performance. We believe that this 'acceptability' study was a crucial part of the research process for developing methods that might routinely be employed within primary care settings. As well as determining the practical acceptability of the instruments, this study has also gained valuable insight into the 'values' which doctors attach to patient assessment of their performance. Despite continued promotion of user and carer involvement, particularly through the mechanism of clinical governance strategies, the caution of participating general practitioners is noteworthy, and lessons should be learned from this. User and carer input into the development of quality within the NHS will only remain as tokenism unless clinicians can be convinced that patients' expectations are rational, and that subsequent assessments are important and meaningful.

This study has highlighted some of the problems that were of most concern to the general practitioners in attempting to measure their performance. Any future attempts to develop measures for routine use in general practice/primary care should take these concerns on board if they are to obtain the co-operation of primary care professionals.

Although the goal is to achieve measures which reflect the core values of their profession and appropriately recognize the burden of their workload, this should not be at the expense of patients' views.

In addition, we have obtained some feedback on general practitioners' views on comparing consultation-based (and patient-based) assessment with output of population-based indicators at practice level. This opened up debate surrounding reward and incentive payments for primary care activities. It was interesting to note that most of these general practitioners would rather see money allocated to help to improve poorly performing practices than to allocate rewards to those who are able to perform better. This appears to be at odds with previous and current attempts to influence quality of care through economic incentives. It may reflect aspiration more than behaviour.

In terms of our research, we found an ambiguous message in that doctors want both measures which are validated against technical competence, and at the same time they see the need to have measures which reflect the doctor–patient relationship and the therapeutic aspect of 'the doctor'. They acknowledge this as the core of general practice. These aspirations could be partially realized by confirmation that enablement is in itself a desirable outcome.

The debates surrounding use of the measures as part of an incentive scheme became focused on the issue of longer consultations (rather than, for example, more skilful consultations). Participants seemed to think that longer consultations would necessarily mean longer consulting hours (which would incur direct costs for doctors), and did not consider that this could be balanced by reduced consulting rates. It is this perceived need for practices to invest even more time in patient care which perhaps influences their call for investment to be put into practices to help them to achieve quality standards, rather than rewarding those who can already deliver. We found no evidence to suggest that practices in less affluent areas had shorter consultation times, but it could be that the doctors within these practices are having to spend more of their time in their practice in order to provide such a level of service. It could also be true that some practices with shorter consultation times spend larger amounts of time on outside commitments, and it would seem inappropriate to pay incentives

to 'poor performers' if this were the case. Clearly we need to understand more about why some practices can provide longer consultations than others before we make any recommendations as to how best to encourage or reward the achievement of this objective. The general practitioners in this study and our own research findings have identified some priorities for future research. Further work is required to determine the appropriateness of the PEI, together with time and continuity measures (Howie *et al.*, 2000), for use among different ethnic groups, and to develop culturally sensitive versions of the instrument as appropriate. We also need to establish whether 'enablement' equates with other improvements in health gain, and this should constitute a wider investigation of whether good interpersonal care can lead to improvements in health and enhance the outcomes of good clinical care. A longitudinal assessment of enablement at consultations should be conducted in order to determine whether 'enablement' is sustained post-consultation.

Acknowledgements

We wish to acknowledge our indebtedness to the 221 doctors and 53 practices who participated, together with their managers and reception staff. Without their tolerance, commitment and goodwill this work could not have been undertaken. We also thank the following, all of whom have contributed to regular discussions about the planning and implementation of this project: Dr T. Jones (Oxfordshire Health Authority), Dr M. Stern (Coventry Health Authority), Dr P. Berrey (Lothian Health Board), Dr R. Elton (University of Edinburgh), Dr M. Pierce (Imperial College School of Medicine) and Dr S. Campbell (National Primary Care Research and Development Centre).

This study was supported by grants from the Chief Scientist's Office at the Scottish Office Home and Health Department, Anglia and Oxford NHS R & D Directorate, West Midlands NHS R & D Directorate and North Thames NHS R & D Directorate.

References

- Baker, R.** 1990: Development of a questionnaire to assess patients' satisfaction with consultations in general practice. *British Journal of General Practice* 40, 487–90.

- Baker, R.** 1991: The reliability and criterion validity of a measure of patients' satisfaction with their general practice. *Fam Practice* 8, 171–77.
- Baker, R.** 2000: Managing quality in primary care: the need for valid information about performance. *Quality in Health Care* 9, 83.
- Barbour, R.S.** 1995: Using focus groups in general practice research. *Family Practice* 12, 328–34.
- Campbell, S., Roland, M., Quale, J., Buetow, S. and Shekelle, P.** 1998: Quality indicators for general practice: which ones can general practitioners and health authority managers agree are important and how useful are they? *Journal of Public Health Medicine* 20, 414–21.
- Campbell, S.M., Roland, M.O., Shekelle, P.G., Cantrill, J.A., Buetow, S.A. and Cragg, D.K.** 1999: The development of review criteria for assessing the quality of management of stable angina, adult asthma and noninsulin-dependent diabetes mellitus in general practice. *Quality in Health Care* 8, 6–15.
- Cantrill, J.A., Sibbald, B. and Buetow, S.** 1998: Indicators of the appropriateness of long-term prescribing in general practice in the United Kingdom: consensus development, face and content validity, feasibility and reliability. *Quality in Health Care* 7, 130–35.
- Cleary, P.D.** 1998: Satisfaction may not suffice: a commentary on 'a patient's perspective'. *International Journal of Technology Assessment in Health Care* 14, 35–37.
- Cleary, P.D. and McNeil, B.J.** 1988: Patient satisfaction as an indicator of quality care. *Inquiry* 25, 25–36.
- Elkan, R. and Robinson, J.** 1998: The use of targets to improve the performance of health care providers: a discussion of government policy. *British Journal General Practice* 48, 1515–18.
- Fitzpatrick, R.** 1991: Surveys of patient satisfaction. I. Important considerations. *British Medical Journal* 302, 887–89.
- Flocke, S.A.** 1997: Measuring attributes of primary care: development of a new instrument. *Journal of Family Practice* 45, 64–74.
- Gelb Safran, D., Kosinski, M., Tarlov, A.R. et al.** 1998: The Primary Care Assessment Survey: tests of data quality and measurement performance. *Medical Care* 36, 728–39.
- Howie, J.G.R., Heaney, D.J., and Maxwell, M.** 1997: *Measuring quality in general practice*. Occasional Paper 75. London: Royal College of General Practitioners.
- Howie, J.G.R., Heaney, D.J., Maxwell, M., Walker, J.J., Freeman, G.K. and Rai, H.A.** 1999: A cross-sectional survey of quality at general practice consultations: principal correlates with enablement as an outcome measure. *British Medical Journal* 319, 738–43.
- Howie, J.G.R., Heaney, D.J., Maxwell, M., Walker, J.J. and Freeman, G.K.** 2000: Developing a 'Consultation Quality Index' (CQI) for use in general practice. *Family Practice* 17, 455–61.
- Kinnersley, P., Stott, N., Peters, T., Harvey, I. and Hackett, P.** 1996: A comparison of methods for measuring patient satisfaction with consultations in primary care. *Family Practice* 13, 41–51.
- Kitzinger, J.** 1995: Introducing focus groups. *British Medical Journal* 311, 299–302.
- Kreuger, R.** 1994: *Focus groups: a practical guide for applied research*. Newbury Park, CA: Sage.
- Levy, S.J.** 1979: Focus group interviewing. In Higginbotham, J.B. and Cox, K.K., editors. *Focus group interviews: a reader*. Chicago: American Marketing Association, 34–42.
- Locker, D. and Hunt, D.** 1978: Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care. *Social Science and Medicine* 12, 283–92.
- McCull, A., Roderick, P., Wilkinson, E., Gabbay, J., Smith, H., Moore, M. and Exworthy, M.** 2000: Clinical governance in primary care groups: the feasibility of deriving evidence-based performance indicators. *Quality in Health Care* 9, 90–97.
- McKee, M.** 1997: Indicators of clinical performance. *British Medical Journal* 315, 142.
- Morgan, D.L.** 1988: *Focus groups as qualitative research*. London: Sage.
- Morgan, D.L.** 1993: *Successful focus groups*. London: Sage.
- Roland, M., Holden, J. and Campbell, S.** 1998: *Quality assessment for general practice: supporting clinical governance in primary care groups*. Manchester: National Primary Care Research and Development Centre.
- Steven, I.D., Thomas, S.A., Eckerman, E., Browning, C. and Dickens, E.** 1999: A patient-determined general practice satisfaction questionnaire. *Australian Family Physician* 28, 342–48.
- Williams, B.** 1994: Patient satisfaction: a valid concept? *Social Science and Medicine* 38, 509–16.
- Wolf, M.H., Putnam, S.M., James, S.A. and Stiles, W.B.** 1978: The Medical Interview Satisfaction Scale: development of a scale to measure patient perceptions of physician behaviour. *Journal of Behavioral Medicine* 1, 391–401.

Box 1 Patient Enablement Instrument

(PEI) completed immediately after the consultation

As a result of your visit to the doctor today, do you feel you are:

	MUCH BETTER	BETTER	SAME OR LESS	NOT APPLICABLE
able to cope with life	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
able to understand your illness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
able to cope with your illness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
able to keep yourself healthy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	MUCH MORE	MORE	SAME OR LESS	NOT APPLICABLE
confident about your health	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
able to help yourself	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Box 2 Focus groups: four main questions

- 1) What do you think of the enablement measure itself? This includes the questions it contains and the responses available.
- 2) What do you think of the data collection process involved?
- 3) What do you think of using the enablement score, time spent at consultations and knowing the doctor well as a measure of the quality of interpersonal care?
- 4) What do you think of the potential of using the quality measure as an incentive-based reward scheme?

Allow approximately 10 minutes of discussion time per question. This is your opportunity to provide feedback on the methods, the instruments and the uses of the results. Any specific questions to the research team should be put forward after the focus group session.