

Do patients and expert doctors agree on the assessment of consultation skills?

A comparison of two patient consultation assessment scales with the video component of the MRCGP

Brian McKinstry, Jeremy Walker^a, David Blaney^b, David Heaney^a and David Begg^c

McKinstry B, Walker J, Blaney D, Heaney D and Begg D. Do patients and expert doctors agree on the assessment of consultation skills? A comparison of two patient consultation assessment scales with the video component of the MRCGP. *Family Practice* 2004; **21**: 75–80.

Objective. The purpose of this study was to determine whether patient ratings of general practice Registrars' consulting skills are associated with 'expert' scoring using the MRCGP video assessment protocol.

Methods. A cross-sectional observational study of general practice Registrars' consultation skills was carried out in 23 practices in South East Scotland using two types of patient assessment compared with expert assessment of video consultation. The main outcome measures were rank correlation of Registrars' overall level of attainment on the Royal College of General Practitioner (RCGP) video assessment with mean score on the Patient Enablement Instrument (PEI) and mean score on the Consultation Satisfaction Questionnaire (CSQ).

Results. The rank correlation of Registrars' mean PEI scores with marks on the RCGP video component was 0.01 ($P = 0.97$, $n = 19$) and mean CSQ score 0.05 ($P = 0.83$, $n = 19$). There were no adverse comments from patients, but Registrars and trainers found the process onerous.

Conclusion. No meaningful association was identified between Registrars' score on the RCGP video examination and patient assessment via either the PEI or the CSQ. This suggests that, with regard to measuring quality in the consultation, one or more of the assessments are invalid or that they are measuring different attributes. Further research to elucidate the reasons for the lack of correlation is required.

Keywords. Assessment, consultation skills, patients, Registrars.

Introduction

Currently, in the UK, the consultation skills of GP Registrars are assessed by examining video portfolios of their consultations by two different systems. One is administered by the Joint Committee on Postgraduate Training in General Practice (JCPTGP) and is designed to assess minimal competence,¹ and the other is the membership examination of the Royal College of General Practitioners (MRCGP).² A smaller number are assessed using simulated surgeries. The video

assessments, although widely accepted, have been criticized by some who believe that the intrusive nature of the assessment affects their content validity, and as yet reliability studies have not been published.^{3–5}

When assessing GP Registrars, patients' views are not formally canvassed or considered. However, a variety of methods of assessing patient satisfaction with the consultation have been described.^{6–8} While the reliability of these measures has been established,^{6–9} studies to determine their criterion validity (does the test measure what it is said to measure?) have not been done, as there is no agreed 'gold standard' against which to judge them.

The use of patient assessment of GP's consultations has become increasingly common. However, it is not clear if quality as measured in this way relates to doctors' views of what constitutes a good consultation. This study sought to determine if there is any meaningful association between patient assessment of consultation skill as measured by the Consultation Satisfaction

Received 3 January 2003; Accepted 8 September 2003.

Lister Institute, Edinburgh, ^aDepartment of Community Health Sciences, University of Edinburgh, ^bAssociate Dean of General Practice, SE Scotland and ^cGP, Penicuik, UK. Correspondence to Brian McKinstry, The Lister, 11 Hill Square, Edinburgh EH8 9DR, UK; E-mail: brian.mckinstry@scpmde.scot.nhs.uk

Questionnaire (CSQ) (see Box 1), consultation outcome as measured by the Patient Enablement Instrument (PEI) (see Box 2) and expert doctor assessment of consultation skills using the video assessment model of the RCGP (see Box 3). The hypothesis of this study was that patient assessments would accurately identify the same ‘good’ consultants as the MRCGP video assessment.

There are several advantages to using general practice Registrars for testing this hypothesis. One practical advantage is that Registrars have to make a video for the summative assessment purposes. In addition, Registrars are a relatively homogenous group of doctors with similar training and experience. They are likely to be equally familiar to the patients they are seeing, a factor known to influence PEI scores.⁷

The purpose of this study was to determine whether patients’ assessments of Registrars’ consulting skills, as measured by the CSQ (which measures satisfaction) and the PEI (which captures a quantity known as ‘enablement’), exhibit significant associations with Registrars’ level of attainment on the video component of the RCGP examination.

Methods

Recruitment

Shortly after beginning the training year, all eligible Registrars ($n = 52$) and their trainers in South East Scotland were sent a letter outlining the project in very general terms. Where both the Registrar and trainer signalled willingness for the Registrar to participate, the practice was contacted to arrange a visit at which the requirements of the study were explained in detail. Those Registrars who agreed to take part were then asked to provide signed consent to the eventual release, under strict confidentiality, of their RCGP video marks to the research team.

Box 1 *The Consultation Satisfaction Questionnaire (CSQ)*

The Consultation Satisfaction Questionnaire (CSQ)⁶ is designed to assess patients’ satisfaction with their general practice consultations. It consists of 18 items, each of which is answered on a 5-point Likert-type scale; the directionality of items is varied to cater for the tendency of some respondents to agree (or disagree) with some statements. The items are shown below. The CSQ attempts to capture four separate dimensions of satisfaction, defined as: ‘general satisfaction’; ‘professional care’; ‘depth of relationship’; and ‘perceived time’. It is a well respected, reliable and frequently used rating scale.

- I am totally satisfied with my visit to this doctor
- This doctor was very careful to check everything when examining me
- I will follow this doctor’s advice because I think he/she is absolutely right
- I felt able to tell this doctor about very personal things
- The time I was able to spend with the doctor was a bit too short
- This doctor told me everything about my treatment
- Some things about my consultation with the doctor could have been better
- There are some things this doctor does not know about me
- This doctor examined me very thoroughly
- I thought this doctor took notice of me as a person
- The time I was allowed to spend with the doctor was not long enough to deal with everything I wanted
- I understand my illness much better after seeing this doctor
- This doctor was interested in me as a person, and not just my illness
- This doctor knows all about me
- I felt this doctor really knew what I was thinking
- I wish it had been possible to spend a little longer with the doctor
- I am not completely satisfied with my visit to the doctor
- I would find it difficult to tell this doctor about some private things

Box 2 *The Patient Enablement Instrument (PEI)*

Patients are asked very shortly before attending their GP to complete forms which request information about the type of problem the patient has come with and his/her expectation of treatment. The patient completes the PEI immediately after the end of the consultation. The PEI is a short (six-item) questionnaire intended to capture the quantity termed ‘enablement’, which is believed to reflect the patient’s ability to understand and cope with his/her illness.

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

	Much better	Better	Same or less	Not applicable
Able to cope with life				
Able to understand your illness				
Able to cope with your illness				
Able to keep yourself healthy				
	Much more	More	Same or less	Not applicable
Confident about you health				
Able to help yourself				

Box 3 *The MRCGP video marking schedule***The competences to be demonstrated**

Results in the video component are issued in the form Fail, Pass or Pass with Merit. Twelve of the performance criteria are preceded by (*P*). These are the criteria which examiners consider to be essential for a result of Pass in consulting skills. A further three performance criteria are preceded by (*M*). These are the criteria which examiners feel must be demonstrated for a result of Pass with Merit in consulting skills. Examiners mark five consultations (each consultation marked by a different examiner), giving the candidate a 'tick' for each competency demonstrated.

Discover the reasons for a patient's attendance

- | | | |
|--------------|-----|--|
| (<i>P</i>) | PC: | the doctor encourages the patient's contribution at appropriate points in the consultation |
| (<i>P</i>) | PC: | the doctor responds to cues |
| (<i>P</i>) | PC: | the doctor elicits appropriate details to place the complaint(s) in a social and psychological context |
| (<i>M</i>) | PC: | the doctor takes the patient's health understanding into account |

Define the clinical problem(s)

- | | | |
|--------------|-----|---|
| (<i>P</i>) | PC: | the doctor obtains sufficient information for no serious condition to be missed |
| (<i>P</i>) | PC: | the doctor chooses an examination which is likely to confirm or disprove hypotheses which could reasonably have been formed OR to address a patient's concern |
| (<i>P</i>) | PC: | the doctor appears to make a clinically appropriated working diagnosis |

Explain the problem(s) to the patient

- | | | |
|--------------|-----|---|
| (<i>P</i>) | PC: | the doctor explains the diagnosis, management and effects of treatment |
| (<i>P</i>) | PC: | the doctor explains in language appropriate to the patient |
| (<i>M</i>) | PC: | the doctor's explanation takes account of some or all of the patient's elicited beliefs |
| (<i>M</i>) | PC: | the doctor seeks to confirm the patient's understanding |

Address the patient's problem(s)

- | | | |
|--------------|-----|--|
| (<i>P</i>) | PC: | the doctor's management plan is appropriate for the working diagnosis, reflecting a good understanding of modern accepted medical practice |
| (<i>P</i>) | PC: | the doctor shares management options with the patient |

Make effective use of the consultation

- | | | |
|--------------|-----|---|
| (<i>P</i>) | PC: | the doctor's prescribing behaviour is appropriate |
| (<i>P</i>) | PC: | the patient and doctor appear to have established a rapport |

Data collection: patient-completed questionnaires

Patients of participating Registrars were, on arrival at the reception desk, given a questionnaire by the receptionist. The original goal was to capture details on 100 consecutive consultations for each Registrar. The first part of the questionnaire was designed to collect basic demographic information about the patient together with a series of questions relating to the nature of the presenting problem(s) in order to determine a reasonable case mix. This portion of the questionnaire was completed by the patient while waiting to see the doctor. On being invited into the consulting room, the patient handed the questionnaire to the Registrar, who recorded the start and end times of the consultation, returning the questionnaire to the patient at the end of the encounter. The patient then returned to the waiting room and completed the final section of the questionnaire, which incorporated two established patient-completed instruments (the PEI and the CSQ). The data were collected during the period late November to January when most Registrars were compiling their videotapes for the MRCGP examination.

Data collection: MRCGP video submissions

For those participating Registrars who (i) made an MRCGP submission and (ii) consented to their results being released to the research team, the marking schedules were provided by the RCGP. The video marking schedule takes the form of a matrix of cells, in which 15 rows represent individual 'Performance Criteria' (PC), and columns represent the five individual markers. The cell content is a 'tick' if the PC in question is deemed, by the individual examiner, to have been met; otherwise, it is blank. Initially, only the first five consultations are assessed—the maximum attainable number of ticks for these consultations is 75 (i.e. 5×15), and the minimum is theoretically zero. Participants' level of attainment on the video examination was evaluated by summing the total number of examiner 'ticks' (possible range: 0–75; observed range: 28–66) awarded. The resulting quantity was used to assess (via rank correlation) the level of association between attainment on the video examination and patient-derived assessments obtained from questionnaires containing the PEI and the CSQ.

Analysis and power calculations

The degree of association between video marks and patient assessments was evaluated via rank correlation. We aimed to recruit 30 Registrars which would have permitted the study to detect as significant (at $P = 0.05$) rank correlations of ≥ 0.36 .

Results

A total of 23 Registrars collected questionnaire data. The total number of questionnaires returned was 2109, with return rates for individual Registrars ranging from 31 to 120. Of the 23 Registrars, RCGP video marking schedules were obtained for 19 participants; the remaining four Registrars did not submit a video tape. Investigation of the associations between (i) mean PEI score and RCGP video performance and (ii) mean CSQ score and video performance was therefore restricted to these 19 participants. As only 19 Registrars provided full data, we did not achieve the degree of power originally envisaged; however, the study had the power to detect as significant (at $P = 0.05$) rank correlation coefficients of ≥ 0.46 . Table 1 shows that the success rates of Registrars in the region participating and not participating in the study was similar.

Summary information describing the questionnaire-derived data for the recruited Registrars is given in Table 2.

The associations of mean PEI score and mean CSQ score with the RCGP video mark were assessed via rank correlation (Spearman's ρ). The rank correlation of the RCGP video mark with mean PEI score was 0.01 [$P = 0.97$; $n = 19$; 95% confidence interval (CI) -0.48 to 0.50]. The rank correlation of the RCGP video mark with mean CSQ score was 0.05 ($P = 0.83$; $n = 19$; 95% CI -0.44 to 0.54). Rank correlations with mean PEI score/mean CSQ score were also calculated for 14 of the 15 individual RCGP video performance criteria—marks for one criterion exhibited zero variability (all candidates achieving the maximum of five examiner 'ticks') so correlations could not be obtained for this criterion. Of the 28 individual correlations calculated (one for each of 14 criteria with mean PEI score, and one for each criterion with mean CSQ score), only one—that of merit performance criterion 11 ('the

TABLE 1 Success rate of participants compared with non-participants

	Passed	Merit	Failed	Total
Taking part in study	13 (68%)	4 (21%)	2 (11%)	19 (100%)
Not taking part in study	24 (65%)	5 (14%)	8 (21%)	37 (100%)

Chi-squared = 1.33; $P = 0.51$ NS

TABLE 2 Summary values of patient-derived measures

Registrar	Mean PEI score	SD PEI	Mean CSQ score	SD CSQ
2	3.5	3.8	68.2	5.5
4	4.7	3.9	69.7	8.9
9	3.4	3.9	66.7	10.4
11	3.3	3.5	66.4	6.0
12	3.1	2.7	68.1	5.1
13	2.3	3.2	64.8	9.8
14	3.5	3.8	68.7	8.6
16	2.8	3.1	65.6	8.2
17	3.8	3.9	69.3	8.7
18	2.9	3.1	67.3	8.5
20	2.9	3.6	66.1	8.6
21	2.5	2.8	66.7	8.3
22	2.4	3.1	66.0	10.3
23	2.9	3.3	63.3	8.9
24	2.8	3.6	68.5	7.4
25	3.3	3.2	67.2	8.3
28	2.6	3.3	64.8	8.7
31	3.9	3.5	65.9	8.3
32	3.0	3.7	66.0	9.1

doctor seeks to confirm the patient's understanding') with mean CSQ score—was significant at the conventional level of 5% ($r_s = 0.50$; $P = 0.03$; $n = 19$; 95% CI 0.06 – 0.78), but failed to reach significance when correlated with the PEI ($r_s = 0.41$; $P = 0.08$; $n = 19$; 95% CI -0.05 to 0.73).

Association of PEI score with CSQ score

Previous work has suggested a significant positive association between scores on the PEI and scores on the CSQ.⁹ In the present study, the rank correlation (Spearman's) of PEI score with CSQ score when aggregated to the level of the individual doctor was 0.62 ($P < 0.01$; $n = 23$; 95% CI 0.28 – 0.82).

Acceptability of the study to Registrars, trainers and patients

The patient data collection was generally well received by patients, and there were very few complaints. Registrars found the process time consuming, although ostensibly their only input to the process was to write the start and finish time on the patient assessment form. They found they often had to explain the process to patients, which they felt took time. Some trainers felt that Registrars were already too stressed with summative assessment and studying for the MRCGP and that it was unfair to ask them to do more.

Discussion

This study suffered difficulties in the recruitment of Registrars and, as a result, there may be some concerns about the representativeness of our sample. However, success rates in the MRCGP did not differ significantly between those who did or did not participate, although it is likely that the more enthusiastic agreed to take part. The smaller number taking part meant that, although correlations between assessment methods were found to be zero, CIs included what might be described as low modest negative or positive correlations.

The MRCGP marking schedule is based on criteria known to be important to patients;^{8,10} therefore, it is reasonable to expect that doctors who display competence in meeting such criteria in a consultation should satisfy patients more than those who do not. The CSQ is said to be a global estimate of satisfaction with the consultation and the PEI is said to be a measure of the effectiveness of consultation, yet the results of the study suggest an absence of any meaningful association between Registrars' consulting skills assessed by expert and patient-based methods. The patient-based assessments were moderately well correlated with one another, suggesting that they were similar in the competencies they were measuring. Other researchers have found at best modest correlations between expert assessment of Registrar consultation skills (using direct observation) and patient assessment using another patient assessment tool (Doctors Interpersonal Skills Questionnaire).¹¹

There are several possible hypotheses to explain these results.

- (i) Patient-based methods and video-based methods measure different competencies.
- (ii) One or both methods fail to measure the stated competencies accurately.
- (iii) Competencies considered important by doctors, e.g. decision sharing, may not be so important to patients.
- (iv) Consultations submitted for video examination are highly selected and not representative of the day-to-day consultations examined by patient assessment. Consequently, they may demonstrate skills which Registrars have acquired but not generalized to other consultations or they may be more challenging than average consultations, and the opportunity to use the behaviours which are scored in the MRCGP video may not occur frequently enough in unselected consultations to influence patient satisfaction or enablement measurably.

Critics of patient assessment argue that the method measures doctors' willingness to comply with patient agendas and the doctor's 'friendliness', while disadvanta-

ging doctors who provide a more challenging and possibly more therapeutic medical encounter.¹² This is evidenced, for example, by reduced 'patient enablement' in patients who expect, but do not receive a prescription during the course of a consultation,¹³ a decision that might be seen as appropriate when assessed by a doctor. There is some evidence that the PEI measures more human qualities such as empathy and respect.¹⁴ Proponents argue that meeting patient agendas and the doctor seeming friendly and approachable are important features of consulting skills and that there is indirect evidence that patients are more likely to comply with treatment when a shared agenda is achieved.¹⁵

There is evidence that while there is considerable agreement between patients and doctors, they do not prioritize consultation competencies in the same way.¹⁶ Some competencies that are considered to be important by consultation experts such as exploring the reasons for the patient's attendance may be interpreted as intrusive by patients, particularly if a psychosocial cause is suspected for a physical complaint. Some patients may not wish to share decision making,¹⁷ resulting in a lower CSQ or PEI score for behaviour considered desirable by experts in consultation behaviour.¹⁵

Tapes made for examination are selected from a large number of consultations. The Registrar, often with the help of their trainer,¹⁸ chooses those that are most likely to demonstrate the required competencies. It may be that these tapes do not reflect the day-to-day consulting of the Registrar⁴ in the way as patient assessment. Possession of a competence is no guarantee of its integration into normal working practice.

Future research should try to determine to what extent videoed consultations reflect day-to-day work, to establish if patients observing videos measure skills in the same way as expert doctors, and to see if the results described above are found with other types of patient consultation assessment and expert doctors assessment (e.g. simulated consultation) and to what extent patients value the same consultation skills as doctors.

Already the General Medical Council is considering the use of patient assessments in revalidation, and it has become part of training practice re-accreditation in the UK. Although this study was performed for pragmatic reasons with Registrars, it is likely that research with more experienced doctors would yield similar results. A question mark must lie over both forms of assessment until further work is done to validate them.

Conclusions

This study showed no meaningful association between video assessment by MRCGP and patients' assessment using either the PEI or the CSQ. Further work is required to identify if the reasons for this result lie fundamentally in the value patients and doctors put on different parts of the consultation or the validity of the assessments themselves.

References

- ¹ Campbell L, Howie J, Murray T. Summative assessment: a pilot project in the West of Scotland. *Br J Gen Pract* 1993; **43**: 430–434.
- ² MRCGP. Video consulting skills module. www.rcgp.org.uk/exam/modules/video/nov00/index.asp
- ³ Bain J, Mackay N. Videotaping of general practice consultations. *Br Med J* 1995; **311**: 952–952.
- ⁴ Howe A. Refusal of videorecording: what factors may influence patient consent? *Fam Pract* 1997; **13**: 233–237.
- ⁵ Coleman T, Manku-Scott T. Comparison of video-recorded consultations with those where patients' consent is withheld. *Bri J Gen Pract* 1998; **48**: 971–974.
- ⁶ Baker R. Development of a questionnaire to assess patients' satisfaction with consultations in general practice. *Br J Gen Pract* 1990; **40**: 487–490.
- ⁷ Howie J, Heaney D, Maxwell M, Walker J, Freeman G, Rai H. Quality at general practice consultations: cross sectional survey. *Br Med J* 1999; **319**: 738–743.
- ⁸ Greco M, Cavanagh M, Brownlea A, McGovern J. The Doctors' Interpersonal Skills questionnaire (DISQ): a validated instrument for use in GP training. *Educ Gen Pract* 1999; **10**: 256–264.
- ⁹ Howie J, Heaney D, Maxwell M, Walker J. A comparison of a patient enablement instrument against two established scales as an outcome measure of primary care consultations. *Fam Pract* 1998; **15**: 165–171.
- ¹⁰ Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: description of the research domain. *Soc Sci Med* 1998; **47**: 1573–1588.
- ¹¹ Greco M, Spike N, Powell R, Brownlea A. Assessing communication skills of GP registrars: a comparison of patient and GP examiner ratings. *Med Educ* 2002; **36**: 366–376.
- ¹² Kremer D. Quality at general practice consultations [letter]. *Br Med J* 2000; **320**: 315.
- ¹³ Howie J, Heaney D, Maxwell M, Walker J, Freeman G. Developing a 'Consultation Quality Index' CQI for use in general practice. *Fam Pract* 2000; **17**: 455–461.
- ¹⁴ Mercer S, Watt G, Reilly D. Empathy is important for enablement [letter]. *Br Med J* 2001; **322**: 865.
- ¹⁵ Tucket D, Boulton M, Olson C, Williams A. *Meetings Between Experts*. London: Tavistock; 1985.
- ¹⁶ Jung H, Wensing M, Grol R. What makes a good general practitioner: do patients and doctors have different views? *Br J Gen Pract* 1997; **47**: 805–809.
- ¹⁷ McKinstry B. Do patients wish to be involved in decision making in the consultation? A cross sectional survey with video vignettes. *Br Med J* 2000; **321**: 867–871.
- ¹⁸ McKinstry B, Dodd M, Blaney D, McLeay G. How much help do trainers give their Registrars in compiling videotapes for summative assessment? A questionnaire survey of trainers and Registrars in two Scottish regions. *Educ Primary Care* 2002; **13**: 55–61.