

# Letters

All letters are subject to editing and may be shortened. Letters should be sent to the BJGP office by e-mail in the first instance, addressed to [journal@rcgp.org.uk](mailto:journal@rcgp.org.uk) (please include your postal address). Alternatively, they may be sent by post as an MS Word or plain text version on CD or DVD. We regret that we cannot notify authors regarding publication. Letters not published in the Journal may be posted online on our Discussion Forum. For instructions please visit: <http://www.rcgp.org.uk/bjgp-discuss>

## Measuring patient satisfaction

In the September issue of *BJGP*, Hankins *et al* published a brief report<sup>1</sup> criticising the reliability and validity of the two survey instruments which are used for measuring patient satisfaction (the (Improving Practice Questionnaire [IPQ] and the General Practice Assessment Questionnaire [GPAQ]) in general practices nationally in order to monitor the quality of services and earn QOF points. I have only been involved in the development of the IPQ and so I will restrict my comments to the reliability and validity of that instrument, and not the GPAQ.

Hankins *et al* conducted a literature search based on the names of the questionnaires IPQ, GPAQ, and GPAS (the forerunner of GPAQ). They only found one paper about the IPQ,<sup>2</sup> none for GPAQ and three papers for GPAS. They also hand-searched journals.

If they had been more thorough, they would have discovered that the IPQ had been previously validated in Australia<sup>3</sup> where it was known as PAIS (Practice Accreditation and Improvement Survey) and that its core consists of the 12 items that comprise another instrument called DISQ (Doctor's Interpersonal Skills Questionnaire)<sup>4-6</sup> which, in spite of its name, can be used to assess any clinicians' capability. There are several papers demonstrating that the DISQ has validity against external criteria, for example against the assessments of College examiners.<sup>7-9</sup>

The IPQ had also been assessed for validity and reliability in 2002 by two independent experts in this field, Professor Mike Pringle and Dr Brian McKinstry, prior to its recommendation as one of the tools to use in the Quality and Outcomes Framework (QOF) of the GP contract.

Criticisms of the IPQ raised by Hankins *et al* were as follows:

- No analysis of reliability was reported.
- Nothing was said about the reliability of the two items that had been found by principal components analysis.
- Correlation of item 27 of the questionnaire with the summed scores for items 1 to 26 does not constitute a measure of validity.
- A significant difference (between the mean satisfaction scores of younger [aged less than 40 years] and older [aged 40 years and over] patients) was interpreted by the authors as evidence of validity.

I still possess the first draft of our paper on the reliability and validity of the IPQ that was submitted to Education for Primary Care in 2003, the final version of which<sup>1</sup> Hankins, *et al* reviewed. In that draft, I described the results of principal components analysis that I had used on the instrument, which yielded two major components that we called 'capacity' (the processes within the surgery) and 'capability' (about the clinician's ability). The capability component incidentally comprised all of the items that make up the DISQ. I had also calculated Cronbach's  $\alpha$  for each of these components. The statistical reviewer of our first draft made cryptic comments about the use of Cronbach's  $\alpha$  that we unfortunately misinterpreted so we regrettably deleted the values from the final draft of the paper. The deleted values were 0.9655 for the capability component and 0.9082 for the capacity component. Values for Cronbach's  $\alpha$  over 0.9 are regarded as demonstrating good internal consistency,<sup>10</sup> and hence reliability. I think this answers the first two criticisms about reliability of the IPQ.

Construct validity is an ongoing process of learning more about a

construct (in this case, satisfaction with general practice), making new predictions and testing them. It is a process where the theory and the process and the measure are assessed at the same time.<sup>11</sup> Satisfaction is a construct because it cannot be directly measured in the way that objective measures such as height or weight can. However, people can tell you how satisfied they feel about something and will be able to give a satisfaction rating. Regarding the validity of the IPQ, we felt that an assessment of construct validity was possible by correlating item 27, which asked patients to rate their overall satisfaction with this visit to their general practice, with the summed values of all the other items (1 to 26). We obtained a correlation coefficient of 0.78,  $P < 0.001$ , which would indicate that the instrument produced responses that suggest it is measuring the construct 'patient satisfaction'. In their report, Hankins *et al* neglected to say what item 27 was and thus detracted from its usefulness in assessing this construct.

Our other approach to evaluating IPQ construct validity was based on comparisons of mean scores across categories such as sociodemographic groups known or hypothesised to score differently. In spite of what Hankins *et al* say, it has generally been the case in past literature (for example, concerning the DISQ) that older patients give better satisfaction scores than younger ones.<sup>10,12</sup> As this was expected, we felt that confirming this to be the case with IPQ demonstrated construct validity.

I first heard about these criticisms after they were presented at a conference in July 2006 and I immediately emailed the lead author to make myself known to him. In retrospect however, he should really have approached us for relevant information before doing his conference

presentation. He replied to my email saying that he had reviewed our paper<sup>1</sup> and asked me if I knew of any further papers. I took that to mean any more recent papers concerning reliability and validity of the IPQ, of which there were none: I didn't realise at the time that he might not have been aware of papers describing the validation of the DISQ. He also asked me to shed more light on the correlation we had done between IPQ items 1–26 against item 27 but I didn't feel there was any more to say about it. He regarded it as an estimate of reliability (and yet in his *BJGP* brief report he said that we hadn't assessed reliability), and I viewed it as a measure of validity. This only serves to underline the fact that validity and reliability are closely related. An unreliable instrument, after all, is unlikely to measure the 'truth'.

Hankins *et al* say in their report that 'access to IPQ and GPAQ datasets may have helped to evaluate reliability' but neither I nor my colleagues received any requests for our IPQ data. We would have gladly given them all the raw data on 55 687 patient responses in 361 UK practices if he had asked for it. It is a shame that they didn't approach us earlier or that the lead author was not more specific in his eventual email reply to me because I could have sent him the Cronbach's  $\alpha$  values for the IPQ and validation references about the DISQ too. In my view, this is not the way to conduct a balanced review. It is regrettable that these matters were not clarified because the Hankins *et al* article remains in the public domain potentially more influential than this complete rebuttal.

### Roy Powell

Coordinator, Peninsula Research & Development Support Unit (RDSU), Exeter.  
E-mail: R.J.Powell@exeter.ac.uk

### REFERENCES

1. Hankins M, Fraser A, Hodson A, *et al*. Measuring patient satisfaction for the Quality and Outcomes Framework. *Br J Gen Pract* 2007; **57**(542): 737–740.
2. Greco M, Powell R, Sweeney K. The Improving Practice Questionnaire (IPQ): a practical tool for general practices seeking patient views. *Educ Prim Care* 2003; **14**(4): 440–448.
3. Greco M, Sweeney K, Brownlea A, McGovern J. The practice accreditation and improvement survey (PAIS) — what patients think. *Aust Fam Physician*

2001; **30**(11): 1096–1100.

4. Greco M, Pocklington S. Incorporating patient feedback into vocational training: an interpersonal skills development exercise for GP trainers and registrars. *Educ Gen Pract* 2001; **12**(3): 285–291.
5. Greco M, Cavanagh M, Brownlea A, McGovern J. Validation studies of the Doctors' Interpersonal Skills Questionnaire. *Educ Gen Pract* 1999; **10**: 256–264.
6. Greco M, Brownlea A, McGovern J. Utilising patient feedback in the RACGP training program: an exploratory study. *Aust Fam Physician* 1995; **24**(6): 1077–1081.
7. Greco M, Spike N, Powell R, Brownlea A. Assessment of GP registrars' communication skills: a comparison of patient and GP examiner ratings. *Med Educ* 2002; **36**: 366–376.
8. Greco M, Brownlea A, McGovern J. Impact of patient feedback on the interpersonal skills of GP registrars: results of a longitudinal study. *Med Educ* 2001; **35**(8): 748–756.
9. Greco M, Francis W, Buckley J, *et al*. Real-patient evaluation of communication skills teaching for GP registrars. *Fam Pract* 1998; **15**(1): 51–57.
10. Bowling A. *Research methods in health: investigating health and health services*. Buckingham: Open University Press, 2002.
11. Streiner DL, Norman GR. *Health measurement scales*. New York: Oxford University Press, 1992.
12. Howie JGR, Porter AM, Heaney DJ, Hopton JL. Long to short consultation ratio: a proxy measure of quality of care for general practice. *Br J Gen Pract* 1991; **41**(343): 48–54.

10.3399/bjgp08X263820

## Authors' response

Roy Powell makes several points in his letter. We will attempt to be brief in our response.

We were aware of the papers he mentions, but the PAIS and the DISQ are different to the IPQ. We found one paper that evaluated the IPQ in a UK sample. We criticised that paper for failing to report the reliability and validity of the IPQ.<sup>1</sup> Powell essentially confirms our four main criticisms of this paper:

- He confirms that the reliability of the IPQ was not reported, a surprising error in a paper that claims to establish the reliability of the IPQ.
- The researchers combined two measures. They then disaggregated them using principal components analysis. This circular logic says nothing about the reliability and validity of the IPQ.
- Older patients had higher scores on the IPQ, but there was no evidence that they were more satisfied. The conclusion that this confirms the validity of the IPQ is,

once again, circular. This also applies to the other papers cited here.

- He suggests that the item-total correlation of  $r = 0.78$  validates the IPQ against the criterion of item 27 of the IPQ. He has this the wrong way around: the item-total correlation validates item 27 against the criterion of the IPQ. In neither case is the IPQ validated against a meaningful criterion.

We have commented on the relevance of the expert panel to this debate.<sup>2</sup> Our review was restricted to peer-reviewed publications. If the independent experts had access to additional data, then perhaps these should be published to allow researchers and other interested parties equal access to the relevant material

We do not agree that we should have approached the IPQ research team before publishing our critical review. We found no peer-reviewed evidence that the QOF patient surveys were valid and reliable and our results have been available since we presented them at conference in 2006.<sup>3</sup> No further relevant data were forthcoming from the IPQ team.

Far from being a 'complete rebuttal', we feel that Powell's comments fully support our findings, and we are pleased to find these issues discussed in a public forum.

### Matthew Hankins

Senior Research Fellow, Division of Primary Care and Public Health, Brighton and Sussex Medical School, Brighton, BN1 9PH.  
E-mail: m.c.hankins@bsms.ac.uk

### Helen Smith

Professor, Brighton and Sussex Medical School

### REFERENCES

1. Greco M, Powell R, Sweeney K. The Improving Practice Questionnaire (IPQ): a practical tool for general practices seeking patient views. *Educ Prim Care* 2003; **14**: 440–448.
2. Hankins M, Smith H. GPAQ. *Br J Gen Pract* 2007; **57**(545): 994.
3. Hankins M, Fraser A, Hodson A, *et al*. Made to Measure? Patient Satisfaction and the Quality and Outcomes Framework. *Fam Pract* 2006; **23**: 61. [http://fampra.oxfordjournals.org/content/vol23/suppl\\_1/index.dtl](http://fampra.oxfordjournals.org/content/vol23/suppl_1/index.dtl) (accessed 10 Dec 2007).

10.3399/bjgp08X263839