

ORIGINAL ARTICLE

The OutPatient Experiences Questionnaire (OPEQ): data quality, reliability, and validity in patients attending 52 Norwegian hospitals

A M Garratt, Ø A Bjærtnes, U Krogstad, P Gulbrandsen

Qual Saf Health Care 2005;14:433–437. doi: 10.1136/qshc.2005.014423

See end of article for authors' affiliations

Correspondence to:
Dr A M Garratt,
Norwegian Knowledge
Centre for the Health
Services, Nasjonalt
Kunnskapssenter for
Helsehjelp, Postboks
7004, 0130 Oslo,
Norway;
andrew.garratt@
kunnskapssenteret.no

Accepted for publication
19 September 2005

Objective: To describe the development and evaluation of the OutPatient Experiences Questionnaire (OPEQ) for somatic outpatients.

Design: Literature review, patient interviews, pretesting of questionnaire items, and a cross sectional survey.

Setting: Postal survey of adult outpatient clinics at 52 hospitals in all five regions of Norway during 2003 and 2004.

Subjects: 35 719 patients who had attended an outpatient clinic within the previous 3 weeks.

Results: 19 266 patients (53.9%) responded to the questionnaire. Low levels of missing data suggest that the questionnaire is acceptable to patients. Factor analysis of items applicable to all patients produced three factors: clinic access (two items), communication (six items), and organisation (four items). The remaining items contributed to the hypothesised scales of hospital standards (three items), information (six items), and pre-visit communication (three items). With the exception of the pre-visit communication scale, the levels of Cronbach's alpha were >0.7. With the exception of the hospital standards scale, all produced test-retest correlations that exceeded 0.7. Most of the results of validity testing were as hypothesised. Correlations between the OPEQ scores ranged from 0.30 (clinic access and hospital standards) to 0.73 (communication and information). As hypothesised, scores were significantly related to patient responses to questions about overall satisfaction, general health and age.

Conclusions: The OPEQ is a self-administered questionnaire that includes the most important aspects of patient experience from an outpatient perspective. It has good evidence for internal consistency, test-retest reliability, and validity.

The measurement of patient perceptions relating to the process and quality of healthcare delivery is increasingly recognised as an important component in the evaluation of healthcare interventions and for assessing service quality.¹ This is reflected in the growth in the use of patient surveys designed to measure concept concepts such as patient satisfaction and patient experiences.^{1, 2}

Patient satisfaction research has been criticised for lacking both a clear definition^{3, 4} and for methodological problems relating to its measurement including validity and reliability.² Satisfaction surveys have traditionally produced very high ratings of patient satisfaction with health care. The reasons for this phenomenon are complex and may include a desire not to appear ungrateful and an acceptance of the limitations of healthcare delivery.⁵ This can lead to a lack of adequate discrimination between good and bad experiences.

An alternative approach involves asking patients to rate their experiences of aspects of health care including communication, information provision, family involvement and the organisation of care.^{6, 7} This form of measurement involves the collection of more objective information relating to whether specific healthcare events occurred—for example, whether they were informed about the results of the examination. There is an inbuilt assumption that the aspects of experience covered by such instruments are related to patient satisfaction. Measures of experience have therefore been referred to as indirect measures of patient satisfaction.¹ It is important that patients are involved in the development of such instruments in order to ensure that the most relevant aspects of the healthcare experience are included.

Over the past 7 years, self-administered postal questionnaires have been used in large scale studies of inpatient experiences with hospital care in the Norwegian healthcare system.^{6, 7} The Patient Experiences Questionnaire (PEQ) for inpatients was found to have evidence for reliability and validity in a sample of 20 890 patients discharged from surgical and internal medicine wards of hospitals across Norway.⁶ The OutPatient Experiences Questionnaire (OPEQ) was developed on the basis of the PEQ with qualitative studies to further inform content validity from an outpatient perspective. The OPEQ was found to have good evidence for reliability and validity in two Norwegian regions.⁸ This study describes the development and evaluation of the OPEQ for somatic outpatients in 52 hospitals across Norway. The evaluation was based on a rigorous process of testing for data quality, reliability, and validity.

METHODS

Development of the questionnaire

The development of the questionnaire followed previous work in the identification of domains and items of relevance to outpatients.^{6–8} The Anglo-American and Scandinavian literature was searched for aspects of patient experiences of importance to an outpatient setting including access, bureaucracy, continuity of care, cost, facilities, humaneness, outcome, overall quality, and psychosocial problems.⁹ As part of a focus group, outpatient clinic staff revised the list of items according to their relevance to Norway.

The questionnaire was piloted through interviews with five patients attending different types of outpatient clinic. The patients completed the questionnaire and were asked to

comment on the relevance of the issues covered and comprehensibility including the response options. Further changes were made following consultation with four physicians and four head nurses from cardiology, gynaecology, neurology, oncology, respiratory medicine, and surgery outpatient clinics. The process of development was designed to ensure content validity—that is, the extent to which the items adequately address important aspects of patient experiences. The 26 items use 10-point scales with descriptors at the end, a response scale which was found to produce a questionnaire with evidence for reliability and validity in Norwegian patients.⁶

Data collection

The questionnaire was mailed to somatic outpatients aged 16 years and over in the Autumn of 2003 and 2004: 12 367 patients from 23 hospitals in the northern and western regions of Norway in 2003 and 23 352 from 29 hospitals in the eastern, middle and southern regions in 2004. Non-respondents were sent a reminder questionnaire after 3 weeks.^{8 10}

Statistical analysis

Data completeness is an indicator of acceptability to patients, so items with high levels of missing data were considered for removal from the questionnaire. Exploratory factor analysis was used to assess the underlying structure of the core items within the questionnaire.¹¹ Factors were extracted with an eigenvalue greater than one. Items with poor factor loadings were considered for removal from the final questionnaire. Following previous findings from research into patient experiences within Norway, it was expected that items would contribute to different aspects of patient experiences including communication, information, and organisation.^{6 8}

Internal consistency was assessed using item-total correlation and Cronbach's alpha. The former measures the strength of association between an item and the remainder of its scale and, following previous findings, it was expected that they would exceed 0.4.^{6 8} The latter assesses the overall correlation between items within a scale. For a scale to be considered sufficiently reliable for use in groups of patients, an alpha value of 0.7 has been recommended.^{12 13}

Test-retest reliability was assessed by sending a second questionnaire to a sample of 270 patients 6 days after they returned the first questionnaire. The former included an additional questionnaire asking patients if they had attended a clinic since the previous questionnaire; this group was not included in the test-retest analysis. Reliability was assessed using the intraclass correlation coefficient which should exceed the criterion of 0.7 for use in groups of patients.¹²

Construct validity was assessed by comparisons of scale scores and responses to additional questions from the postal survey. Following the derivation of scale scores supported by the preceding analyses, it was hypothesised that the largest correlation above 0.7 would be found between the communication and information scales. These scales were expected to produce moderate correlations between 0.5 and 0.7 with the organisation scale. Correlations with other aspects of patient experiences assessed by the questionnaire were expected to fall below 0.5.

It was hypothesised that scale scores would correlate with responses to single items assessing overall satisfaction,¹⁴ perceived treatment correctness, and the organisation of examinations and tests. Correlations above 0.5 were hypothesised for scales relating to communication, information, and organisation. It was hypothesised that patients reporting their health as poor would have significantly lower scores than those reporting their health as good,¹ and that scores would have small levels of correlation below 0.4 with

responses to a question relating to the perceived effect of the visit on the patient's health problem.¹⁵ The strongest associations were expected for the communication and information scale scores.

Two hypotheses related to appointment times: (1) that patients who wished to change their appointment and found it easy to do so would have higher scores than those who did not; and (2) that patients who had their appointment changed without being consulted would have lower scores than those who did not. The largest differences were expected for the pre-visit communication scale. It was also hypothesised that patients attending a follow up visit who saw the same clinician would have higher scores than those who saw a different clinician. The largest differences were expected for the three scales most closely related to the clinician encounter: communication, information, and organisation. Finally, it was hypothesised that scores would have a small correlation with age.¹

RESULTS

Data collection

Of the 35 719 patients who were mailed a questionnaire, 19 266 (53.9%) responded. The mean (SD) ages of respondents and non-respondents were 55.5 (17.4) years and 52.1 (20.5) years, respectively. Compared with respondents, non-respondents were more likely to be male (40.8% v 46.7%). These differences were statistically significant.

Statistical analysis

The levels of missing data, responses to the "does not apply" category, and descriptive statistics are shown in table 1. Missing data ranged from 1.7% to 5.9%. It was decided that items with the "does not apply" option could not constitute the core questionnaire as they were not relevant to a large proportion of patients. The item relating to whether there was enough time for dialogue was retained because it had a low rate of "does not apply" responses and is potentially an important aspect of patient experiences relating to communication.

As has been widely documented in the literature, mean item scores are skewed towards positive experiences.^{9 14} The lowest and highest mean scores are for the acceptability of the appointment waiting time and cleanliness items, respectively.

Factor analysis produced three factors which accounted for 64.9% of the total variation between patients (table 2). They can be described as clinic access, communication, and organisation. The factor loadings were acceptable.

The levels of item-total correlation for the core questionnaire items are acceptable, ranging from 0.54 to 0.73 (table 1). The alpha values meet the criterion of 0.7, ranging from 0.76 to 0.85 for clinic access and communication, respectively. Table 1 also shows the hypothesised scales for the remaining items that were not applicable to a large proportion of patients. These scales were based on the literature review which informed the development of the questionnaire and include hospital standards, information, and pre-visit communication. Pre-visit communication has lower item-total correlations and Cronbach's alpha, which meets the less stringent criterion of 0.5.¹⁶ Items within hospital standards and information are sufficiently correlated with the remainder of the scale scores and produce acceptable levels of alpha.

Of the 270 patients mailed a test-retest questionnaire, 194 (71.9%) responded and, of these, 148 did not have another clinic visit. Five of the scales produced reliability estimates above 0.8 and, with the exception of hospital standards, all exceeded the criterion of 0.7 (table 1).

The results of validity testing are shown in tables 3 and 4. The correlations between the communication, information,

Table 1 Descriptive statistics, internal consistency, and test-retest reliability

Scale/item	Missing* (%)	Does not apply (%)†	Mean‡	SD	Cronbach's alpha/ item-total correlation (n = 18829)	Test-retest intraclass correlation (n = 139)
Clinic access	329 (1.7)		89.67	17.48	<i>0.76</i> [§]	0.81
Ease of finding clinic/ward	388 (2.0)		8.06	1.77	0.62	
Ease of finding way within clinic/ward	382 (2.0)		8.13	1.64	0.62	
Communication	278 (1.4)		85.32	19.15	<i>0.85</i>	0.81
Enough time for dialogue	357 (1.9)	1582 (8.2)	7.53	2.37	0.65	
Person understandable	321 (1.7)		7.88	2.18	0.54	
Person competent	326 (1.7)		8.05	1.94	0.62	
Person caring	334 (1.7)		7.71	2.04	0.73	
Opportunity to give sufficient information	450 (2.3)		7.58	2.36	0.73	
Unanswered questions	702 (3.6)		7.17	2.83	0.61	
Organisation	476 (2.5)		85.55	18.25	<i>0.80</i>	0.80
Background information available	573 (3.0)		7.82	2.16	0.59	
Staff collaboration	524 (2.7)		8.01	1.77	0.68	
Organisation of work	518 (2.7)		7.70	1.84	0.66	
Person well prepared	451 (2.3)		7.51	2.34	0.55	
Hospital standards	5754 (29.9)	1562 (8.1)	89.72	15.83	<i>0.75</i>	0.65
Waiting room	1131 (5.9)	1365 (7.1)	7.72	2.04	0.54	
Toilet	886 (4.6)	3439 (17.8)	8.12	1.78	0.64	
Cleanliness	1005 (5.2)	1793 (9.3)	8.33	1.40	0.60	
Information	3744 (19.4)	2534 (13.1)	80.72	23.43	<i>0.88</i>	0.87
Information-self-care	334 (1.7)	4409 (22.9)	7.62	2.43	0.69	
Information-medication/side-effects	566 (2.9)	12795 (66.4)	6.83	3.04	0.69	
Information-examinations	595 (3.1)	3702 (19.2)	7.94	1.98	0.69	
Information-examination/test results	624 (3.2)	2666 (13.8)	7.36	2.63	0.69	
Information-condition/prognosis	572 (3.0)	3141 (16.3)	6.55	3.14	0.72	
Consulted about examination/treatment	513 (4.7)	7500 (38.9)	7.16	2.88	0.64	
Pre-visit communication	4646 (23.6)	3808 (19.8)	82.82	21.77	<i>0.59</i>	0.84
Acceptability of appointment waiting time	836 (4.3)	7805 (40.5)	6.54	2.87	0.41	
Information from clinic received in advance	518 (2.7)	7390 (38.3)	7.37	2.70	0.43	
Ease of accessing clinic staff	442 (2.3)	8884 (46.1)	7.44	2.36	0.37	

*For scale scores the number of missing values is the sum of responses to the "does not apply" category and true missing values—that is, where a scale score is not calculable.

†For scale scores the figure for "does not apply" is equal to the number of patients for whom a scale score was not calculable because they responded "does not apply" to half or more items within the scale.

‡Items are scored 0–9 and scales are scored 0–100 where higher scores represent better experiences.

§Values for Cronbach's alpha are in italics.

and organisation scores range from 0.59 to 0.73, the largest being for communication and information (table 3). The correlations between these and the pre-visit communication scale range from 0.47 to 0.51. The remainder of the correlations range from 0.31 (clinic access and information) to 0.44 (hospital standards and organisation).

The majority of the six scale scores have moderate to large correlations with overall satisfaction, ranging from 0.33 to

0.69 for clinic access and communication, respectively. Patient perceptions of the correctness of treatment have small to moderate correlations ranging from 0.27 to 0.53 for clinic access and communication, respectively. Responses to the question relating to the organisation of tests or examinations have the largest correlation with organisation. The perceived effect of the clinic visit on the health problem has small levels of correlation with scale scores, the largest being for communication and information. Following previous findings,¹ age is positively correlated with patient experiences. Finally, the reported waiting time has small negative correlations with several of the scores, the largest being for pre-visit communication, communication, and organisation.

Table 4 shows further results of the validity testing. As has been widely documented,¹ compared with patients in better health, those in poor health had significantly poorer experiences on four of the scales. For patients who had to change appointment, those who found it easy to do so had significantly higher scores, the differences being largest for pre-visit communication. The six scores were also significantly lower for patients who had their appointment moved or changed without asking. Again, the score differences were largest for pre-visit communication. Finally, for patients attending a follow up visit, those seeing the same clinician had significantly higher scores, the difference being largest for communication and information.

Table 2 Factor analysis with loadings (n = 16 022)

Item	Factor*		
	1	2	3
Ease of finding clinic/ward			0.78
Ease of finding way within clinic/ward			0.79
Background information about patient available		0.58	
Staff collaboration		0.87	
Organisation of work		0.74	
Person well prepared		0.46	
Enough time for dialogue	0.62		
Person understandable	0.55		
Person competent	0.53		
Person caring	0.68		
Opportunity to give sufficient information	0.93		
Unanswered questions	0.65		
Variation%	45.10	11.93	7.88

*Factor loadings above 0.3 are reported.

Table 3 Correlations between scales scores and responses to individual questions (n = 19 000)

Scale/variable	Clinic access	Communication	Hospital standards	Information	Organisation	Pre-visit communication
Communication	0.35					
Hospital standards	0.38	0.38				
Information	0.31	0.73	0.34			
Organisation	0.39	0.66	0.44	0.59		
Pre-visit communication	0.33	0.48	0.33	0.47	0.51	
Overall satisfaction	0.33	0.69	0.39	0.63	0.63	0.46
Perceived treatment correctness	0.27	0.53	0.35	0.51	0.49	0.36
Organisation tests/examinations	0.37	0.50	0.43	0.45	0.62	0.44
Effect of visit on health problem	0.05	0.31	0.06	0.37	0.22	0.23
Age	0.19	0.24	0.24	0.17	0.27	0.20
Waiting time	-0.10	-0.16	-0.09	-0.15	-0.16	-0.23

Data shown as Spearman's rank correlations (*r*).
All correlations are significant ($p < 0.01$).

DISCUSSION

The OPEQ is a short self-completed questionnaire that is acceptable to patients while maintaining comprehensibility in its coverage of important aspects of patient experience.⁸ Questionnaire development was based on an extensive literature review and the views of patients and clinicians who felt that the relevant aspects of patient experiences were adequately covered. The questionnaire comprises three core scales that are widely applicable to outpatients: clinic access, communication, and organisation. The three remaining scales of hospital standards, information, and pre-visit communication are not applicable to all patients and should be assessed for relevance within clinical specialties and at different organisational levels before application. The generic core scales can be supplemented by these and other aspects of patient experiences of relevance to specific patient groups. Further involvement of patients in this process will help ensure that the questionnaire has content validity as a measure of patient experiences.

The instrument has undergone a rigorous process of testing for reliability and validity, which support its application as a measure of patient experiences. The core scales are supported by the results of the factor analysis. The high levels of "does not apply" responses meant that it was not possible to include the remaining items in the factor analysis, but high levels of internal consistency reliability for information and hospital standards suggest that the items comprising these hypothesised scales are sufficiently related. With the exception of pre-visit communication, they have good levels of internal consistency reliability. The pre-visit communication scale meets the less stringent reliability criterion of 0.5¹⁶ and

has good test-retest reliability. With the exception of hospital standards, the remainder of the scales also produced test-retest estimates in excess of 0.8. Hospital standards fell just below the criterion of 0.7.

The OPEQ has good evidence for construct validity with the hypotheses largely being met. In the comparison of scale scores, those measuring related aspects of experience including communication and information had the highest levels of correlation. The significant relationships between scale scores and age, health, and overall satisfaction follows previous findings.¹⁻¹⁴ The majority (22/24) of the group comparisons followed the direction hypothesised and were statistically significant. Compared with those who rated their health as good, patients who rated their health as poor had significantly lower scores for four of the six scales. Patients who rated their health as poor had slightly higher scores for clinic access. These patients are more frequent attendees, which is a plausible explanation for this finding. Finally, for patients attending a follow up visit, seeing the same clinician significantly improved their experiences, the largest differences being for the communication and information scales.

Given the large sample sizes, it is not surprising that some of these differences were statistically significant. However, while some of the differences are quite small, the majority (15/24) are in excess of five points and five are over 10 points. Because of the general finding that patient satisfaction scores are usually skewed towards the more positive end of the spectrum,⁹⁻¹⁴ these differences are potentially important and, if found at the ward or organisational level, should be considered for investigation. Moreover, the majority of the larger differences relate to communication, information and

Table 4 Mean (SD) scale scores for variables with hypothesised associations

Variable/scale	Clinic access	Communication	Hospital standards	Information	Organisation	Pre-visit communication
Poor health						
No (n = 16554)	89.61 (17.35)	86.11 (18.34)	89.73 (15.66)NS	82.01 (22.37)	85.94 (17.82)	83.52 (21.13)
Yes (n = 2209)	90.14 (18.34)	79.29 (23.53)	89.40 (17.26)	71.57 (28.26)	82.71 (20.97)	77.85 (25.30)
Easy to change appointment						
Yes (n = 7676)	90.99 (16.53)	87.45 (17.21)	90.66 (15.13)	83.40 (21.59)	87.95 (16.33)	86.31 (18.79)
No (n = 1167)	85.66 (20.78)	75.95 (25.77)	85.34 (20.41)	71.46 (28.26)	77.15 (24.42)	69.19 (28.00)
Appointment changed						
Yes (n = 4293)	88.49 (18.35)	83.51 (20.08)	88.97 (16.30)	78.83 (23.72)	83.72 (19.43)	79.55 (23.92)
No (n = 11878)	90.30 (16.91)	86.23 (18.55)	90.14 (15.51)	81.78 (23.00)	86.66 (17.38)	84.60 (20.35)
Same clinician at follow up						
Yes (n = 8325)	91.79 (15.42)	88.34 (16.54)	90.51 (15.18)	84.72 (20.34)	88.89 (15.09)	86.14 (19.53)
No (n = 4851)	89.86 (17.38)	80.90 (21.91)	88.86 (16.57)	75.33 (26.03)	81.01 (21.53)	80.18 (23.32)

All differences are statistically significant ($p < 0.01$) unless indicated (Mann-Whitney test).
NS, not significant.

Key messages

- The measurement of patient experiences is an important component in the evaluation of healthcare delivery.
- The OutPatient Experiences Questionnaire (OPEQ) is based on reviews of the literature and the views of patients and health professionals.
- The OPEQ has good evidence for internal reliability, test-retest reliability, and validity.
- The OPEQ is an appropriate measure of patient experiences for outpatient clinics across Norway.

organisation, and pre-visit communication which were hypothesised. The association between patient experiences including pre-visit communication and the conduct of appointments is an important finding. Keeping to original appointment times and allowing patients to change appointments where necessary may improve experiences with outpatient care in Norway.

The low response rate found by this study is cause for concern. It is below the mean response rate based on a systematic review,¹ but similar to those found in previous surveys of patient experiences in Scandinavia.^{6,17} Non-respondents are more likely to be members of minority groups and less well educated.¹ The present study found that non-respondents were more likely to be younger and male, which was also found in a study of psychiatric outpatients attending clinics in Norway.¹⁸ The study findings also support the large body of evidence that older patients tend to report higher levels of satisfaction.¹ Non-respondents might therefore have had poorer experiences, but this requires further research.

In summary, the OPEQ is acceptable to patients and has good evidence for data quality, internal and test-retest reliability, and validity. The instrument is recommended in future applications designed to assess patient experiences of outpatient clinics. It is being used to measure patient experiences in hospitals throughout Norway.

ACKNOWLEDGEMENTS

The authors thank Kjell Ingar Pettersen for his valuable comments, Tomislav Dimoski for information technology support, and Saga Høgheim, Nina Viksløkken Ødegård and Reidun Skårerhøgda for help with data collection. The views expressed are those of the authors.

Authors' affiliations

A M Garratt, Ø A Bjærnes, U Krogstad, Norwegian Knowledge Centre for the Health Services, Nasjonalt Kunnskapssenter for Helsestjenesten, Oslo, Norway

P Gulbrandsen, Medical Faculty Division, Akershus University Hospital, University of Oslo, Oslo, Norway

This research was funded by the Norwegian Social and Health Directorate (SHDir).

Competing interests: none.

The Norwegian Regional Committee for Medical Research Ethics, the Data Inspectorate and the Norwegian Board of Health approved the survey.

REFERENCES

- 1 **Crow R**, Gage H, Hampson S, *et al.* The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002;**6**(32).
- 2 **Sitzia J**. How valid and reliable are patient satisfaction data? An analysis of 195 studies. *Int J Qual Health Care* 1999;**11**:319–28.
- 3 **Avis M**, Bond M, Arthur A. Satisfying solutions? A review of some unresolved issues in the measurement of patient satisfaction. *J Adv Nurs* 1995;**22**:316–22.
- 4 **Sitzia J**, Wood N. Patient satisfaction: a review of the issues and concepts. *Soc Sci Med* 1997;**45**:1829–43.
- 5 **Fitzpatrick R**. Capturing what matters to patients when they evaluate their hospital care. *Qual Saf Health Care* 2002;**11**:306.
- 6 **Pettersen KI**, Veenstra M, Guldvog B, *et al.* The Patient Experiences Questionnaire: development, validity and reliability. *Int J Qual Health Care* 2004;**16**:453–63.
- 7 **Guldvog B**, Hofoss D, Pettersen KI, *et al.* PS-RESKVA (Patient Satisfaction, Results and Quality)-patient satisfaction in hospitals. *Tidsskr Nor Laegeforen* 1998;**118**:386–91.
- 8 **Garratt AM**, Andresen Ø, Krogstad U, *et al.* Pasienterfaringsinstrumentet PasOpp i somatiske poliklinikker. *Tidsskrift Nor Laegeforen* 2005;**125**:421–4.
- 9 **Hall JA**, Dorman MC. What patients like about their medical care and how often they are asked: a meta-analysis of the satisfaction literature. *Soc Sci Med* 1988;**27**:935–9.
- 10 **Andresen Ø**, Garratt A, Groøven G, *et al.* Pasienterfaringer ved somatiske poliklinikker i Helse Øst, Sør og Midt-Norge – Metodedokumentasjon, PasOpp rapport nr 6-2004. Oslo: Nasjonalt Kunnskapssenter for Helsestjenesten, 2004.
- 11 **Hair J**, Anderson RE, Tatham RL, *et al.* Factor analysis. In: *Multivariate data analysis with readings*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- 12 **Nunnally JC**, Bernstein IH. *Psychometric theory*, 3rd ed. New York: McGraw-Hill, 1994.
- 13 **Kline P**. *A handbook of test construction*. London: Methuen, 1986.
- 14 **Jenkinson C**, Coulter A, Bruster S, *et al.* Patients' experiences and satisfaction with health care: results of a questionnaire study of specific aspects of care. *Qual Saf Health Care* 2002;**11**:335–9.
- 15 **Carmel S**. Satisfaction with hospitalization: a comparative analysis of three types of services. *Soc Sci Med* 1985;**51**:1817–25.
- 16 **Helmstadter GC**. *Principles of psychological measurement*. New York: Appleton-Century-Crofts, 1964.
- 17 **Jenkinson C**, Coulter A, Gyll R, *et al.* Measuring the experiences of health care for patients with musculoskeletal disorders (MSD): development of the Picker MSD questionnaire. *Scand J Caring Sci* 2002;**16**:329–33.
- 18 **Garratt AM**, Bjørngård JH, Dahle KA, *et al.* The Psychiatric Out-Patient Experiences Questionnaire (POPEQ): development, reliability and validity. *Nordic J Psychiatry* (in press).