



COVER SHEET

Whitehouse, Sarah L and Blom, Ashley W and Taylor, Adrian H and Pattison, Giles T R and Bannister, Gordon C (2005) The Oxford Knee Score; Problems and Pitfalls. *The Knee* 12(4):pp. 287-291.

Accessed from <http://eprints.qut.edu.au>

Copyright 2005 Elsevier

Author copy:

Whitehouse SL, Blom AW, Taylor AH, Pattison GTR, Bannister, GC. The Oxford Knee Score; problems and pitfalls:- *The Knee* 2005;12:287-91

The Oxford Knee Score; Problems and Pitfalls

Sarah L Whitehouse
PhD

Adrian H Taylor[†]
FRCS

Ashley W Blom
FRCS

Giles TR Pattison
FRCS

Gordon C Bannister

[†] Corresponding author
Mr AH Taylor
Department of Orthopaedic Surgery
University of Bristol
Winford Unit
Avon Orthopaedic Centre
Westbury on Trym
Bristol BS10 5NB

Tel: 0117 959 5906

Fax: 0117 959 5924

e-mail: adtaylor50@hotmail.com

ABSTRACT:

The Oxford Knee Score is a self-completed patient based outcome score. We audited the outcome of total knee arthroplasty at our unit using the Oxford Knee Score. The hypothesis of this study is that the OKS can be easily and accurately completed by unassisted patients.

Of 856 patients who had undergone total knee arthroplasty and were given questionnaires, 769 (90%) responded. 624 (81%) of the respondents managed to complete the questionnaire. A number of the 12 items composing the questionnaire posed problems for the patients and a number of items were left blank. Item 4 (concerning walking time) was omitted in 82 (13%) of the 624 completed questionnaires. Calculation of Cronbach's alpha for internal consistency suggests that there are redundancies within the Score.

Limitations in some of the items of the scale suggest the need for reconsideration and reformulation of questions and response categories. This study suggests that where detailed assessment of outcome is required, such as for outcome studies or controlled trials, the Oxford Knee Score, in its present form, is not ideal for use as a postal questionnaire.

KEYWORDS: outcome, arthroplasty, replacement

RUNNING HEAD: The Oxford Knee Score; Problems and Pitfalls

INTRODUCTION

As the number of instruments available to assess the effectiveness of interventions grows, it becomes increasingly important to ensure the careful and complete validation of these tools. This is important whether the measure is patient or surgeon based, and as the emphasis is increasingly on the involvement of the patient in this assessment, these tools need diligent evaluation. The potential for postal rather than clinical assessment has enormous economic implications.

The Oxford Knee Score (1) (OKS) was a natural follow-on from the Oxford Hip Score (2) and the Oxford Shoulder Score (3). It is a 12 item questionnaire for patients having total knee replacement and was developed from patient interview and validated against two generic health scales, the SF-36 (4) and Health Assessment Questionnaire (HAQ) (5). It was originally intended for use in large randomised controlled trials for patients undergoing total knee replacement, to assess levels of, and changes in, pain and function of the knee solely from the patients viewpoint. Data collection can be via a postal questionnaire or by directed interview.

The OKS was designed to be short, practical, reliable, valid and sensitive to clinically important change, hence being more accurate than other patient based measures, such as the SF-36 or the Arthritis Impact Measurement Scale (AIMS) (6). These other scales have been criticised for their length, difficulty in completion, lack of responsiveness and relevance for joint replacement patients (7). When compared with the Knee Society Score (8), SF-36 and HAQ, the OKS fared favourably in terms of reproducibility, internal consistency, validity and responsiveness. It was therefore selected as the principal outcome tool in a postal audit of patients following total knee replacement from the authors' institution with at least two years follow-up.

However, since the commencement of this project, problems have been encountered with the scale. Patients experienced difficulty completing the score, raising concerns about the clarity and validity of the scale. This has been similar to experiences encountered with the Oxford Hip Score (9).

Although assessing the validity and clarity of the scale were not the aims of the audit, the problems encountered with it require comment and further investigation.

MATERIALS AND METHODS

The Oxford Knee Score (Table I) was incorporated into a postal audit of all patients who received total knee replacement at one purpose built, regional, elective orthopaedic centre. Follow-up was at least two years. In addition to the OKS, patients were questioned regarding complications encountered postoperatively, including readmission and presence and extent of infection. Pain maps were completed by those patients with persistent pain.

Each item of the Oxford Knee Score has allocated to it a score; 1 for the least limited response (e.g. no pain) and 5 for the most severe (e.g. severe pain) as described in Table I. Scores are then calculated by adding item scores together to give an overall score of between 12 (no pain or limitation for all items) to 60 (severe pain or limitation for all items).

Only those patients having a single, primary knee were included in the analysis.

The study population consisted of patients receiving a single unilateral total knee replacement between April 1993 and August 1996 at the Avon Orthopaedic Centre, Bristol, under the care of 18 Consultant Orthopaedic Surgeons. Patients were contacted by letter including a questionnaire. Failure to respond to this initial mailing within 6 months prompted a further letter. Those not responding to this were located through Family Health Service Authority database records to check for change of address. Patients known to be still alive were either contacted by telephone or sent a further request (if their address had changed).

RESULTS

Demographics: There were 956 unilateral, primary total knee replacements carried out between April 1993 and August 1996. One hundred patients had

died, and of the remainder, 769 primary cases responded, giving an overall follow-up rate of 90.8% (868 / 956).

Five hundred and three (52.6%) patients had a right sided operation, and 644 (67.4%) patients were female. The demographics of those patients responding were similar: 405 (52.7%) were right sided, 520 (67.6%) were female. The responders are highly likely therefore to be representative of the population as a whole.

Of the 769 cases who completed the questionnaire, 685 (89.1%) responded to the postal questionnaires, while the remainder were contacted by telephone interview.

Infection rates: Ninety five (12.4%) of the 769 patients had been readmitted to hospital for a complication following surgery. These included thromboembolic events, infection, manipulation under anaesthesia (for stiffness), intensive physiotherapy, arthroscopy, wound healing problems and medical problems.

One hundred and seven (13.9%) patients reported more specific problems with infection; 39 requiring extra antibiotics whilst in hospital and 31 having a delay in discharge because of infection (17 both), 55 had a reddened or inflamed wound after discharge requiring additional antibiotics from their General Practitioner, 29 required readmission for further investigation or treatment and 13 (1.7%) had a deep infection (inside the joint) requiring further surgery.

There were also 185 patients (24.1%) who reported persistent pain since the operation, a value significantly higher than the response to the corresponding question following primary total hip replacement (17.4% reported persistent pain for this cohort of 1377 patients).

Missing item and clarity of item responses: Of the 769 patients responding to the questionnaire, 624 (81.1%) gave complete Oxford Knee Scores. In 13 (1.7%), all items were missing.

Figure I indicates the number of missing responses for each item. In a few cases, the whole of one side of the questionnaire (which is two-sided) was not completed.

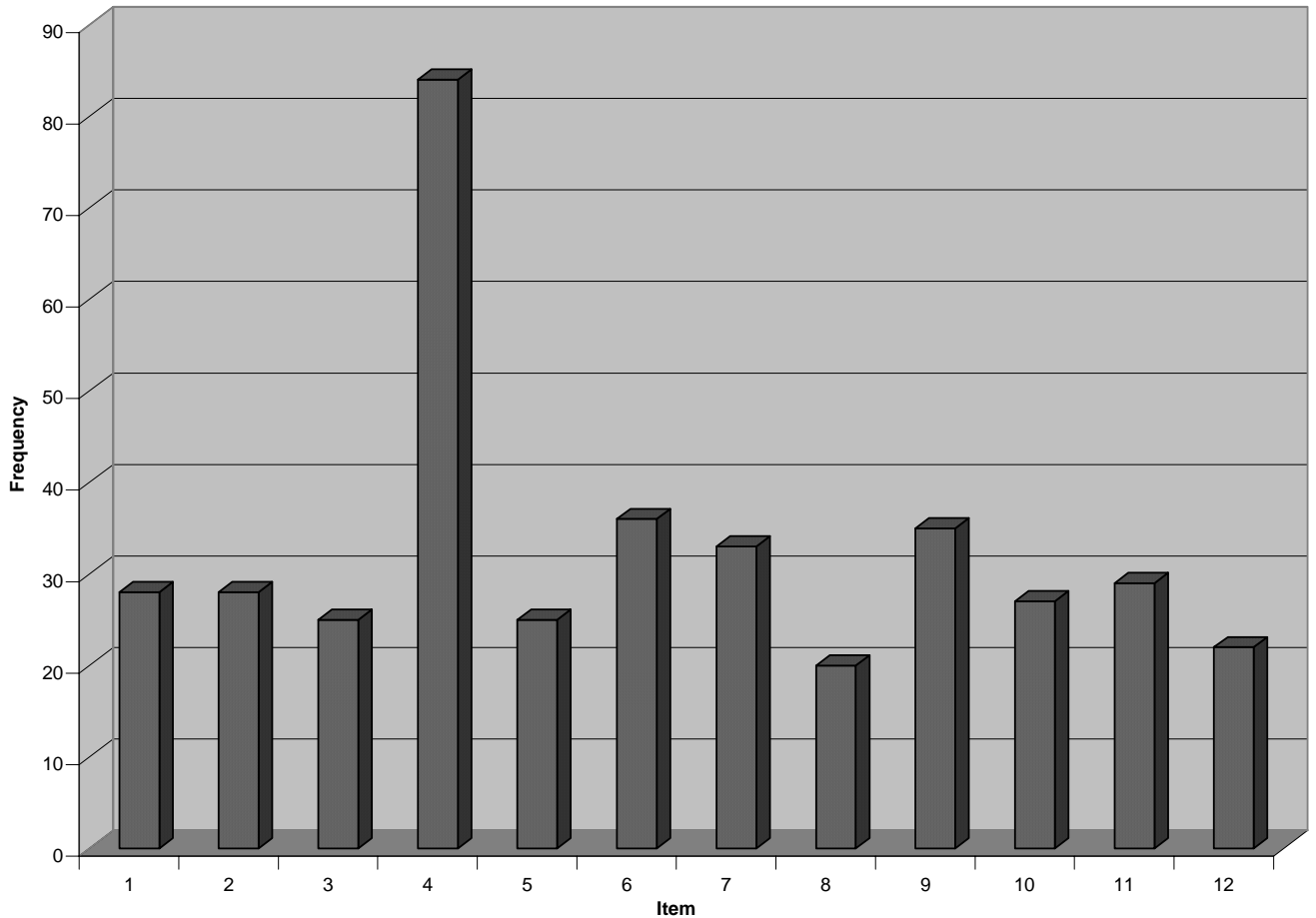


Figure I. Frequency of missing responses per item of the Oxford Knee Score

Item 4 - Walking time: The significantly higher frequency of missing responses for item 4: *For how long have you been able to walk before pain from your knee becomes severe? (with or without a stick)* can be explained by the lack of clarity of the response categories accompanying the item. This is in accordance with McMurray (9) and the findings of a parallel study for total hip replacement patients performed at our centre (10). Confusion arose for many patients around the first response category - *no pain/more than 30 minutes* - many patients thought that this did not amount to the same thing. There also appeared to be a lack of distinction between the first and the last response category (*not at all – pain severe on walking*). On some occasions, the item was selected and the *pain severe on walking* part crossed through,

thus indicating that this response was selected in error, when the first was the correct response. More often, both responses were selected. Most common were additional written comments from the patients who felt the need to clarify their responses. This was not limited to this item, however, and comments were sometimes made for many of the items. Indeed some patients composed an accompanying letter in order to fully clarify their overall responses.

This lack of clarity in the construct of the response categories also resulted in patients selecting the last response, rather than the first, without providing any other comment. Although this is impossible to detect with confidence without patient interview, it may be deduced by examining the questionnaires of those patients who indicated responses of 1's and 2's for the rest of the OKS, and a 5 for this particular item. This phenomenon occurred in 22 (3.5%). This calculation probably underestimates this error rate.

Item 7 – Kneeling: Special attention should be drawn to this particular item; *could you kneel down and get up afterwards?* Once again, many patients commented on the inappropriateness of the item, as many were told specifically not to attempt this particular activity. Patients therefore selected the last response (*no, impossible*), even when scoring well in all other items. In addition, patients reported pain or discomfort after first attempting to kneel, and hence only knelt in emergency situations, which led to more punitive scoring. When all other responses were 1 or 2, this item had response 5 in 23 (3.7%) cases. Again, this is probably an underestimate of the misclassification of this item.

Many patients commented on this item, attempting to clarify their response with comments such as “advised not to”, “do not kneel”, “due to other joints” and “never tried”.

The inclusion of this particular item makes it very difficult for patients to attain a ‘perfect’ score of 12 points, even when they are pain free and not functionally limited in any other way. Of the 57 patients indicating a score of

1 for all other items, only 19 (33.3%) selected response 1 for the kneeling item.

Comorbidities: Difficulties also arose where patients commented that they had more than one condition affecting their overall ability to complete the questionnaire. Comorbidities are well established as confounders to outcome after joint replacement and must therefore be treated with care. Some patients commented that they were unable to distinguish between problems arising from their index knee as opposed to the contra-lateral knee or arthritis of the hip. Likewise some patients had difficulty performing general tasks because of medical problems rather than orthopaedic problems, and as such were limited by this co-existent condition rather than by their index knee operation. It then becomes impossible to speculate as to the effect that the knee has on these functions and becomes invalid. This problem is further confounded if the presence of comorbidities is not part of the score and their influence cannot be accounted for.

Patients undergoing total knee replacement are generally elderly and it is rare to find patients without either comorbid medical conditions or arthritis affecting other joints.

Oxford Knee Scores: The score is generally reported on 12 (best) to 60 (worst) scale. Often this is misconstrued, as the majority of other scoring systems are reported on a 0 (worst) to 100 (best) scale. In the OKS the direction is inverted and the scale is distributed around a mid-value of 36 points. There are no published population normals for the scale, and so what constitutes a good score is based purely on clinical experience.

For the 624 valid responses for this audit, the median score was 26.5 points with an inter-quartile range of 19 to 36 points (range 12 to 57).

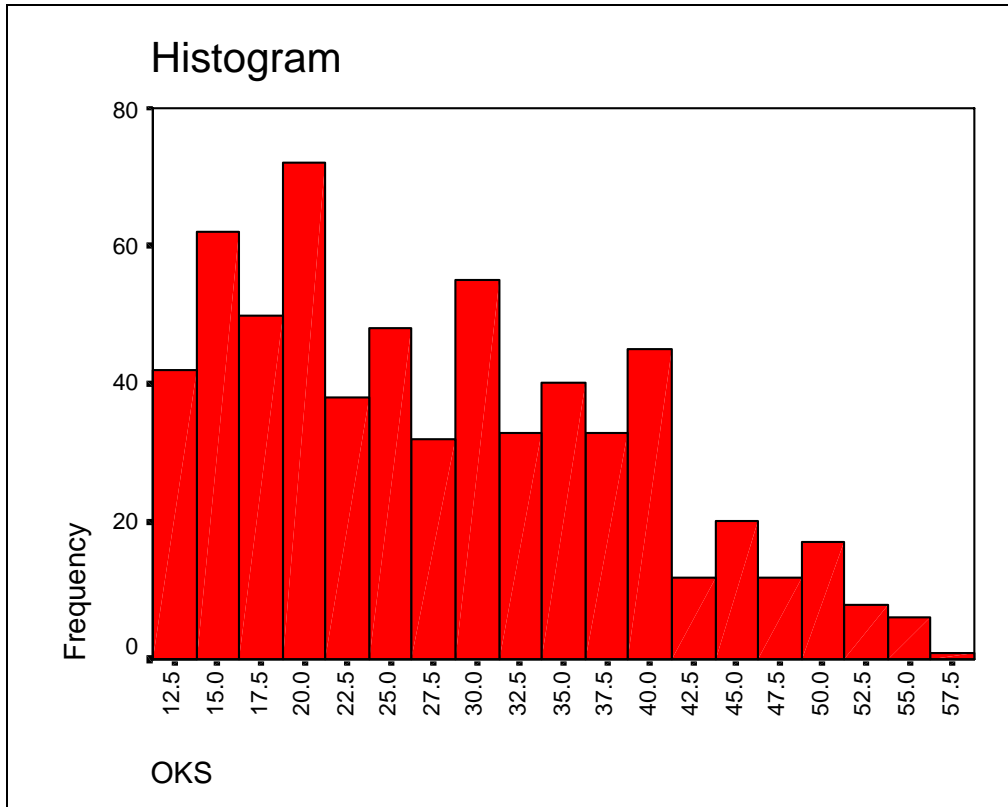


Figure II. Histogram of Oxford Knee Scores

Figure II indicates the skewness of the distribution. The majority of respondents are towards the left of the graph, indicating a more favourable score.

Internal consistency: The internal consistency of a scale is a function of the number of items and their covariation. Random error due to item selection is modelled in this estimate of reliability. The coefficient commonly used to estimate the reliability of the instruments based on internal consistency is Cronbach's alpha (11). It is calculated using a two-way fixed effects model, which measures the agreement between items. The reliability of the average of the items is generally of interest. Values of Cronbach's alpha of between 0.7 and 0.9 show adequate reliability for a scale, whereas values above this indicate redundancies in the scale (12).

```

***** Method 1 (space saver) will be used for this analysis ****
R E L I A B I L I T Y   A N A L Y S I S   -   S C A L E   ( A L P H A )

Item-total Statistics

Item                Scale      Scale      Corrected      Alpha
                   Mean       Variance   Item-          if Item
                   if Item   if Item    Total          Deleted
                   Deleted   Deleted   Correlation
1                  25.5641    98.4100    .7631          .9097
2                  26.1250   104.8704    .6803          .9140
3                  25.5625   103.3733    .7659          .9111
4                  25.7532   107.4735    .3732          .9278
5                  25.8429   101.4263    .6424          .9152
6                  25.7564    99.6613    .7017          .9125
7                  24.1026   103.3410    .5630          .9187
8                  25.8397   100.9791    .7015          .9125
9                  25.7244    99.1053    .8442          .9068
10                 26.2292   105.6472    .6745          .9145
11                 25.4279    97.1055    .7132          .9123
12                 25.4022   100.1509    .7595          .9101

Reliability Coefficients

N of Cases =      624.0                N of Items = 12

Alpha =      .9205

```

Figure III. SPSS output for reliability analysis of OKS

The only item which, if removed, increases the value of Cronbach's alpha is item 4; *for how long have you been able to walk before the pain from your knee becomes severe? (with or without a stick)*, which is one of the items which presented problems for respondents. The smallest fall in Cronbach's alpha occurs when item 7; *could you kneel down and get up again afterwards?* another item posing problems for patients, is deleted from the model. When both these items are deleted from the model, the valid number of cases rises to 683 (an increase of 59 cases) and Cronbach's alpha rises to 0.9289.

However, this high value of alpha suggests that there may well be redundancies within the scale, and in particular with regards to those items already discussed.

CONCLUSIONS

The ambiguity of the response categories for item 4 and the inappropriateness of the kneeling item (item 7) make the acceptance of this scale for the assessment of outcome after total knee replacement uncertain. Information regarding comorbidities is not included in the scale and this militates against its use as a sole outcome measure.

The scale, as for the Oxford Hip Score, lacks the scope and sophistication necessary to accommodate the complex multiple and inter-related nature of many patients' problems (McMurray 1999). The intentional oversimplification of this important feature of the patients' condition, makes completion by some patients neither possible nor valid.

Since the use of this scale demands concurrent use of other instruments in order to attain sufficient information renders the fact that the scale is short and easy to complete, becomes less relevant.

DISCUSSION

This study was commenced as an audit project to assess, by means of a postal questionnaire, patients with a minimum two years follow-up after total knee replacement. It was not intended as a validation study of the Oxford Knee Score, but has rather highlighted several problems in the administration of this instrument. Had validation of the instrument been one of the aims of the study, then concurrent completion of well established instruments to assess aspects of health related quality of life would have been essential.

Patients experience difficulty completing the Oxford Knee Score, and its use should therefore be carefully monitored. Its use as a postal questionnaire may be limited as accurate completion cannot be guaranteed.

The OKS does provide a simple and brief scale for the assessment of outcome after total knee replacement and these scores are quick and easy to calculate and analyse. Where scores are also collected pre-operatively, progress can be readily assessed.

Limitations in some of the items of the scale suggest the need for reconsideration and reformulation of questions and response categories. It is unclear as to the precise role the questionnaire plays in the assessment of outcome. This study suggests that where detailed assessment of outcome is required, such as for outcome studies or controlled trials, the Oxford Knee Score, in its present form, is not suitable.

Where a brief summarised account of outcome is sufficient, however, such as for audit purposes, the Oxford Knee Score is an acceptable instrument.

TABLES

Item	Response categories
<i>During the past four weeks</i>	
1) How would you describe the pain you <u>usually</u> have from your knee?	1 None 2 Very mild 3 Mild 4 Moderate 5 Severe
2) Have you had any trouble washing and drying yourself (all over) <u>because of your knee</u> ?	1 No trouble at all 2 Very little trouble 3 Moderate trouble 4 Extreme difficulty 5 Impossible to do
3) Have you had any trouble getting in and out of a car or using public transport <u>because of your knee</u> ? (<i>whichever you tend to use</i>)	1 No trouble at all 2 Very little trouble 3 Moderate trouble 4 Extreme difficulty 5 Impossible to do
4) For how long have you been able to walk before the <u>pain from your knee</u> becomes severe? (<i>with or without a stick</i>)	1 No pain/>30 min 2 16 to 30 min 3 5 to 15 min 4 Around the house only 5 Not at all – severe on walking
5) After a meal (sat at a table), how painful has it been for you to stand up from a chair <u>because of your knee</u> ?	1 Not at all painful 2 Slightly painful 3 Moderately painful 4 Very painful 5 Unbearable
6) Have you been limping when walking, <u>because of your knee</u> ?	1 Rarely/never 2 Sometimes or just at first 3 Often, not just at first 4 Most of the time 5 All of the time
7) <i>Could</i> you kneel down and get up again afterwards?	1 Yes, easily 2 With little difficulty 3 With moderate difficulty 4 With extreme difficulty 5 No, impossible
8) Have you been troubled by <u>pain from your knee</u> in bed at night?	1 No nights 2 Only 1 or 2 nights 3 Some nights 4 Most nights 5 Every night
9) How much has <u>pain from your knee</u> interfered with your usual work/housework?	1 Not at all 2 A little bit 3 Moderately 4 Greatly 5 Totally

10) Have you felt that your knee might suddenly 'give way' or let you down?	1 Rarely/never 2 Sometimes or just at first 3 Often, not just at first 4 Most of the time 5 All of the time
11) Could you do the household shopping <u>on your own</u> ?	1 Yes, easily 2 With little difficulty 3 With moderate difficulty 4 With extreme difficulty 5 No, impossible
12) Could you walk down a flight of stairs?	1 Yes, easily 2 With little difficulty 3 With moderate difficulty 4 With extreme difficulty 5 No, impossible

Table I. The Oxford Knee Score

REFERENCES

- 1 Dawson J, Fitzpatrick R, Murray D, Carr A: Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg [Br]* 80-B(1):63-9, 1998
- 2 Dawson J, Fitzpatrick R, Carr A: Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg [Br]* 78-B:185-90, 1996
- 3 Dawson J, Fitzpatrick R, Carr A: Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg [Br]* 78(4)-B:593-600, 1996
- 4 Ware JE, Snow K, Kosinski M, Gandek B. SF-36 health survey manual and interpretation guide. Boston MA. The Health Institute, New England Medical Center, 1993
- 5 Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: The Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 9:789-93, 1982
- 6 Meenan R, Gertman P, Mason J. Measuring health status in arthritis: The Arthritis Impact Measurement Scales. *Arthritis Rheum* 1980;23:146-52
- 7 Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D: Quality of life measure in health care. I: applications and issues in assessment. *BMJ* 305:1074-7, 1992
- 8 Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop* 1989;248:13-4
- 9 McMurray R, Heaton J, Sloper P, Nettleton S: Measurement of patient perceptions of pain and disability in relation to total hip replacement: the place of the Oxford hip score in mixed methods. *Quality in Health Care* 8:228-33, 1999

- 10 Taylor AH, Pattison G, Whitehouse SL, Blom A, Bannister GC. Postal questionnaires – do they work? JBJs [Br] 83-B supp 1:79, 2000
- 11 Cronbach LJ: Coefficient alpha and the internal structure of tests. Psychometrika 16:297-334, 1951
- 12 Streiner DL, Norman GR: Health measurement scales. A practical guide to their development and use. Second edition. Oxford University Press, 2000