Quality of Life Research (2006)
DOI 10.1007/s11136-006-0061-5

A patient-based questionnaire to assess outcomes of foot surgery: Validation in the context of surgery for hallux valgus

Jill Dawson[1], Jane Coffey[2], Helen Doll[1], Grahame Lavis[3], Paul Cooke[3], Mark Herron[4] & Crispin Jenkinson[1]
[1]*Department of Public Health, University of Oxford, Old Road, Oxford, UK (E-mail: jill.dawson@dphpc.ox.ac.uk);* [2]*School of Health and Social Care, Oxford Brookes University, Marston Road Campus, Jack Straws lane, Oxford, UK;* [3]*Nuffield Orthopaedic Centre, Windmill Road, Oxford, UK;* [4]*The Royal Orthopaedic Hospital, The Woodlands, Bristol Road South, Birmingham, UK*

Accepted in revised form 7 March 2006

Abstract

Background: A patient-based outcome measure with good measurement properties is urgently needed for use in clinical trials of foot surgery. *Methods*: We evaluated an existing foot pain and disability questionnaire (the Manchester Foot Pain and Disability Questionnaire) for its suitability as an outcome measure in the context of hallux valgus corrective surgery. Interviews with patients led to initial changes, resulting in 20 candidate questionnaire items with five response categories each. These were tested in a prospective study of 100 patients (representing 138 foot operations) undergoing hallux valgus corrective surgery. Analysis of underlying factor structure, dimensionality, internal reliability, construct validity and responsiveness of the questionnaire items in relation to (i) SF-36 general health survey and (ii) American Orthopaedic Foot & Ankle Society (AOFAS) hallux clinical scale resulted in a final 16 item questionnaire (the 'Manchester-Oxford Foot Questionnaire' (MOXFQ)), consisting of three domains/scales: 'Walking/standing' (seven items), 'Pain' (five items) and 'Social interaction' (four items) each having good measurement properties. All three domains were unidimensional. *Conclusions*: The new 16-item MOXFQ has good measurement properties in the context of outcomes assessment of surgery for hallux valgus. Future studies should assess the MOXFQ in the context of surgery for other foot and ankle conditions.

Key words: Foot, Function, Hallux valgus, Pain, Patient-based outcome measure, Surgery

Introduction

Foot and ankle surgery constitutes around 15–20% of orthopaedic practice [1]. However, many forms of foot and ankle surgery have not been adequately evaluated due to the lack of appropriate, standardised methods of outcomes assessment – and more specifically, those that are patient-focused [2, 3]. Patient assessed outcomes are key in this area because foot problems impair many aspects of health-related quality of life [4].

While a number of questionnaires have been produced with the intention of measuring patients' subjective experience of foot problems, these have generally not been developed with patients' input (eg. Foot Function Index [5], Foot Health Status Questionnaire [6]; Musculoskeletal Outcomes Data Evaluation and Management System (MODEMS) AAOS outcomes questionnaire [7]).

The measurement properties of outcome measures also need to be evaluated in the appropriate context. Currently, no patient-generated measure

of foot problems has been validated as an outcome measure to evaluate foot surgery, or in relation to a specific foot problem – such as hallux valgus.

We report on the development and assessment of a short questionnaire intended for use as an outcome measure of foot surgery. The measure was adapted and tested in the context of patients about to undergo surgery for hallux valgus. We concentrated on this condition, in the first instance, because it is common – up to a third of the population is affected by the deformity [8, 9] – and clinical trials of surgical interventions for the condition are urgently needed. This is because during the last hundred years, around 150 different surgical procedures have been described [2], while studies report between a quarter and a third of patients to be dissatisfied with the outcome [10].

This paper reports the developmental stages, item selection and assessment of the baseline measurement properties, together with the responsiveness of this new foot outcome measure.

Materials and methods

Local ethics committee approval was obtained (Applied and Qualitative Research Ethics Committee reference A02.009) and all subjects consented to participate in the study.

Development of the instrument

A literature review revealed no evidence of appropriately validated patient generated outcome measures for foot surgery. However, one measure: the Manchester Foot Pain and Disability Questionnaire (MFPDQ) [11], appeared potentially suitable, as it had been developed to assess foot pain and function using interviews with people that included patients with foot problems attending clinical rheumatology and podiatric settings. We therefore began with an assessment of the MFPDQ questionnaire items for their appropriateness within a surgical context.

The MFPDQ comprises 19 items (see Table 1), all prefaced with the words: 'Because of pain in my foot', with three possible responses offered for each question ('none of the time', 'on some days', 'on most/every day'). Its development and measurement properties have previously been assessed within the context of a large cross-sectional population survey (n = 1000) of men and women, with questionnaire responses also obtained from 45 rheumatology patients and 33 patients who had attended their GP with a foot problem [11]. Within that study, factor analysis had revealed it to have four underlying sub-scales – one representing function, two pain and one concerned with personal appearance.

Interviews

In order to test the suitability and the content validity of the MFPDQ as an outcome measure for use within a surgical context, the current study began with exploratory interviews conducted with 10 patients who were attending hospital surgical out-patient clinics for hallux valgus.

Interviews were semi-structured and employed prompts regarding patients' perceived problems – pain, mobility (including sports activities), footwear, effects on work/domestic work, social life – associated with their foot condition, plus a general exploration of their feelings about their foot/feet. Lasting around 20 min, interviews ended by asking patients to complete and comment on individual items from the MFPDQ.

Interviews continued with patients until no new common themes emerged. These revealed that, within the surgical context, while the main themes covered by the MFPDQ remained relevant, there were no items concerned directly with pain severity and the number and wording of response categories appeared overly restrictive and, at times, ambiguous. A number of changes were therefore made to the questionnaire with items added, removed, and amended (see Table 1). The response categories were increased from 3 to 5 per item, scored on a Likert-type scale from 0 (representing no limitation) to 4 (representing maximum limitation). The resulting list of 20 candidate questionnaire items, together with the precise wording of response options, is shown in Table 1. These were then piloted on a further seven patients. No further changes were indicated at this stage.

Study to test the questionnaire

A prospective study, to test this candidate questionnaire (to be called 'the Manchester-Oxford

Table 1. Details of the original MFPDQ items and changes leading to the first (20-item) version of the MOXFQ

Original Manchester questionnaire items [11]	Final wording of candidate items, following interviews, tested in longitudinal study (including new items "+")	Items removed ("X") following analysis
Because of pain in my feet	*During the past 4 weeks, this has applied to me*	
1. I avoid walking outside at all	1. I avoid walking outside because of pain in my right/left[a] foot	**X**
2. I avoid walking long distances	2. I avoid walking long distances because of pain in my right/left foot	
3. I don't walk in a normal way	3. I change the way I walk due to pain in my right/left foot	
4. I walk slowly	4. I walk slowly because of pain in my right/left foot	
5. I have to stop and rest my feet	5. I have to stop and rest my right/left foot because of pain	
6. I avoid hard or rough surfaces when possible	6. I avoid some hard or rough surfaces because of pain in my right/left foot	
7. I avoid standing for a long time	7. I avoid standing for a long time because of pain in my right/left foot	
8. I catch the bus or use the car more often	8. I catch the bus or use the car in stead of walking, because of pain in my right/left foot	
9. I need help with housework/shopping	9. I need help with housework or shopping because of pain in my right/left foot	**X**
10. I get irritable when my feet hurt	10. I get irritable due to pain in my right/left foot	**X**
11. I feel self-conscious about my feet	11. I feel self-conscious about my right/left foot	
12. I get self-conscious about the shoes I have to wear	12. I feel self-conscious about the shoes I have to wear	
13. I have constant pain in my feet	13. I have pain in my right/left foot	
14. My feet are worse in the morning	14. The pain in my right/left foot is worse in the morning	**X**
15. My feet are more painful in the evening	15. The pain in my right/left foot is worse in the evening	
16. I get shooting pains in my feet	16. I get shooting pains in my right/left foot	
17. I am unable to carry out my previous work	17. The pain in my right/left foot prevents me from carrying out my work/everyday activities	
18. I no longer do all my previous activities (sport, dancing, hill-walking etc)	18. I am unable to do all my social or recreational activities because of pain in my right/left foot	
	19. *During the past 4 weeks*, how would you describe the pain you usually have in your right/left foot? +	
	20. *During the past 4 weeks*, have you been troubled by pain from your right/left foot in bed at night? +	
(item 10 in original) I still do everything but with more pain or discomfort	(Item removed following interviews)	
Response categories (all items)	*Response categories*	
None of the time	*Items 1–18 inclusive*	
On some days	None of the time, rarely, some of the time	
On most/every day(s)	Most of the time, all of the time	
	Item 19	
	None, very mild, mild, moderate, severe	
	Item 20	
	No nights, only 1 or 2 nights, some nights	
	Most nights, every night	

[a]Questionnaire items are foot-specific and we obtained separate responses for right and/or left foot – according to which foot was to undergo treatment.

Foot Questionnaire' (MOXFQ)), was conducted on patients who underwent surgery. The baseline, pre-operative stage assessed patients within 4 weeks prior to surgical correction for hallux valgus, with follow-up assessments conducted at 12 months post-surgery. The period of 12 months was recommended by the study's clinicians, as we wished to allow a sufficient time, following surgery, to permit full recovery (including getting back into sports activities – where applicable) to

have occurred for the majority of patients. This would allow us to assess the comparative ability of different instruments to measure change in peoples' health status.

Sample sizes for questionnaire development are typically based upon the assumption that the number of respondents should exceed the number of items in a questionnaire by at least a factor of three [12]. So, with around 20 candidate items, a sample size of 100 was considered sufficient.

Between August 2002 and February 2004, 111 consecutive patients were sent a letter of invitation within 2 weeks prior to receiving foot surgery. Of these, 100 consented to take part in the study, 38 of whom were booked to receive bilateral hallux valgus correction (= 138 'cases' i.e. each foot = one case).

Assessments

Patients completed a questionnaire, unassisted, while attending a pre-admission clinic. This comprised (1) demographic items, (2) the MOXFQ (containing 20 items at this stage), with higher scores denoting greater severity, completed separately for each foot to be operated on; and (3) the SF-36 general health questionnaire [13, 14]. The SF-36 contains 36 items and is widely used as a generic health status instrument. It provides scores on eight dimensions: physical functioning, social functioning, role limitations due to physical problems, role limitations due to emotional problems, mental health, energy/vitality, bodily pain and general health perceptions over the last 4 weeks. Scores for each dimension range from 0 (poor health) to 100 (good health). An extra item addresses health change during the last 12 months. Summary scores can be calculated from the eight dimensions to represent limitations related to physical (physical component summary, PCS) and mental (mental component summary, MCS) factors [15] which are standardised to general population scales with a mean of 50 (SD 10).

A foot surgeon or consultant surgical podiatrist completed a clinical examination using the standard American Orthopaedic Foot & Ankle Society (AOFAS) hallux metatarsophalangeal (MTP)-interphalangeal (IP) clinical scale [16]. The AOFAS hallux scale has not been formally validated and its measurement properties may be sub-optimal [17]. Nevertheless, it has been widely adopted by clinicians and there can be little doubt, from the wording of items, that it has face validity and is clearly focused on measurements concerning the foot (and the hallux specifically). It therefore measures what it purports to measure, sufficient to support an *a priori* hypothesis that at least a moderate correlation would be obtained with the new measure (MOXFQ) which purports to measure something similar.

In this scale, a maximum score of 100 points is possible in a patient with no pain, full range of MTP and IP motion, no MTP or IP instability, good alignment, no limitation of daily or recreational activities, and no footwear limitations. Forty points are assigned to pain, 45 to function and 15 to alignment. The surgeon was blinded to patients' responses to the other health status instruments.

The same assessments were repeated, in an outpatient clinic setting at 12 months post-surgery.

Statistical methods

Data were analysed primarily using SPSS version 11.5 [18]. Non-parametric tests were used in the analysis where data were found to be non-normally distributed. To facilitate a general population comparison, SF-36 scores were adjusted for age and sex. Data are presented using mean with standard deviation (SD), median with range, or N (%), as appropriate.

Examination of the dimensionality of the MOXFQ, and the functioning and fit of individual items, was undertaken using a Rasch unidimensional measurement model in RUMM2010 [19–21]. Since 38 patients had bilateral operations, (which threatened the independence of patients' observations on their two individual feet), baseline analyses were repeated three times: on the data for left feet only (n = 65), right feet only (n = 73), and for both feet (n = 138; with 38 patients contributing data for both left and right feet). Similar checks were undertaken with the analysis of responsiveness. The results were, in fact, very similar for all analyses, and thus only the analyses that combined data for left and right feet (n = 138 booked operations, n = 126 post-surgery) are presented here. The significance level throughout was set at two-sided $p < 0.05$.

Item exclusion

Items were excluded from the MOXFQ on the basis of two pre-determined criteria: (1) if, in terms of their response distribution, they showed high ceiling or floor effects, or (2) if, on fitting to a Rasch unidimensional model to any identified domains, they showed particularly poor fit to the model. Items were also considered for revision or exclusion if, on factor analysis, they cross-loaded on more than one factor.

Floor/ceiling effects

A floor or ceiling effect was considered to exist where at least 50% of responses to an item took either of the two most extreme response categories.

Factor structure

Exploratory factor analysis with principal components extraction and orthogonal (varimax) rotation was performed to examine the factor structure of the MOXFQ. Factors were extracted if their eigenvalue was >1. Domain scores of the resulting factors were calculated as the sum of the component item scores.

Individual item functioning

The MOXFQ and its component dimensions were assessed using the one parameter Rasch model [22, 23]. This assumes that as a person's disability or symptoms increase, the probability of a maximum score on the item increases. The Rasch model assesses the unidimensionality of items in a scale. While the total items in the MOXFQ were not expected to fit a unidimensional model, any domains identified by factor analysis were expected to fit such a model and thus to confirm the structure of the MOXFQ.

Internal consistency

Cronbach's alpha coefficients were calculated to assess the internal consistency of the MOXFQ domains. Values of alpha in the range 0.80–0.90 are considered optimal [24], with a minimum alpha of 0.70 being necessary to claim internal consistency [25, 26].

Convergent validity

The convergent validity of the MOXFQ was assessed using Spearman's correlation coefficients between any individual MOXFQ domain scores that emerged (e.g. pain) and similar domain scores of the SF-36 and the AOFAS hallux clinical scale. Similar domains from different instruments were expected to be highly correlated ($r>0.5$) with each other. Thus, a MOXFQ pain domain would be expected to be highly correlated with the SF-36 pain domain; a MOXFQ domain concerned with walking function domain would be expected to be highly correlated with the SF-36 physical functioning and role physical domains and with the SF-36 PCS score.

Divergent validity

The divergent validity of the MOXFQ was assessed using Spearman's correlation coefficients between MOXFQ domain scores and dissimilar individual domain scores on the SF-36 and the AOFAS hallux clinical scale. Thus, a MOXFQ pain domain would not be expected to be highly correlated with the SF-36 Role emotional, Social functioning, Mental health, Energy/vitality and General health perception domains or with the SF-36 MCS score.

Responsiveness

Responsiveness was assessed by comparing mean pre- and 12 month post-operative scores, change scores and effect sizes between MOXFQ, AOFAS hallux clinical scale and SF-36 domains on the cases that received surgery ($n=126$ foot operations) where complete data were provided at both pre-and post-operative assessments. Change scores were calculated as the pre-operative score minus the post-operative score for each instrument. Effect size is a method of calculating the extent of change measured by an instrument in a standardised way that allows comparison between instruments [27]. Here, it was calculated as the difference between the sample's mean pre-surgical and post-surgical scores, divided by the SD of pre-surgical scores. An effect size of 1.0 is equivalent to a change of one SD in the sample. Effect sizes of 0.2, 0.5 and 0.8 are typically regarded as indicating small, medium and large degrees of change respectively.

It was hypothesised that at least moderate effect sizes (≥ 0.5) would be obtained from the foot-specific measures and relevant domains of the SF-36 (i.e. pain and physical functioning) and that the

foot specific measures would be the most responsive.

Results

Study sample and characteristics

The baseline study sample consisted of 100 people representing 138 ft (i.e. each foot = one 'case') booked for surgery. At baseline, the mean age of study sample patients (n = 100) was 50.03 (SD 12.87, median 52, range 20–75) years. Almost all (95%) were female, less than half (41/100, 41%) were in full-time paid employment, with 27 (27%) part-time employed and 21 (21/100, 21%) retired (the remaining 12/100, 12% were students, unemployed or homemaker/carer). No significant differences were found regarding demographic characteristics when people booked for bilateral vs. unilateral surgery were compared. Of the 62 patients due to have unilateral surgery, 35 (56%) were having surgery on their right foot. Nine people (13 foot operations) had their surgery cancelled or postponed.

Pre-operative scores for clinical assessment and SF-36

At baseline, the patients' mean scores on the AOFAS hallux clinical scale were as follows: *right foot*: mean 54.8 (SD 13.51) minimum 22.0, maximum 83.0; *left foot*: mean 55.5 (SD 11.0) minimum 27.0, maximum 81.0. Patients' scores on the SF-36 are given in Table 2. In comparison with population norms [28] patients in the study had significantly lower (poorer) pain, physical functioning, role physical, and general health perception scores.

Item response distribution

None of the 20 MOXFQ items showed a ceiling effect, but two items showed a particularly strong floor effect. Thus, to item 9 ('help with housework'), 77% responded 'none of the time' and to item 1 ('avoid walking outside'), 51% responded 'none of the time' and 29% 'rarely'. A decision was taken to exclude these two items – since they would not be responsive to improvement – and to mark another two items (item 14 'pain worse in morning' and item 10 'I get irritable') for potential exclusion.

Initial assessment of dimensionality and factor structure of the MOXFQ

Dimensionality

The remaining 18 items were found to be non-unidimensional following Rasch analysis.

Factor structure

An exploratory factor analysis of the remaining 18 items extracted three factors with an eigenvalue >1, explaining 63.2% of the variance. However, two items (item 14: 'pain worse in the morning' and item 10: 'I get irritable') cross-loaded on two factors and, as these items had also shown particularly skewed responses, it was decided to exclude these items from the MOXFQ.

Table 2. Pre-operative SF36 scores (mean (SD)) in study sample, in a general population sample [28], and in the study sample adjusted for age and sex

SF36 domain	Study sample (n = 100) mean (SD)	Study sample adjusted for age and sex mean	General population [28] (n = 9332) mean (SD)	Significance test (*T*-test) *p* =
Physical functioning	76.74 (21.63)	78.82	88.40 (17.98)	< 0.001
Pain	62.01 (24.55)	56.89	81.49 (21.69)	< 0.001
Role physical	77.21 (27.04)	69.47	85.82 (29.93)	< 0.001
Role emotional	82.91 (24.23)	77.09	82.93 (31.76)	0.085
Social functioning	79.90 (23.21)	76.61	88.01 (19.58)	0.063
Mental health	70.35 (16.77)	67.38	73.77 (17.24)	0.188
Energy/vitality	57.32 (19.72)	56.61	61.13 (19.67)	0.549
General health perception	76.94 (18.09)	72.72	73.52 (19.90)	0.002

The MOXFQ 16-item questionnaire (final version – see Appendix 1)

Completion rates

The completion rate for the 16 individual items was good and the majority of items only attracted 1 or 2 missing responses (maximum 4). Scale totals (see below) were only computed for individuals where responses were obtained for all items comprising that scale. Completion rates for the three domains were as follows: Walking/standing 135/138 (97.8%); Pain 133/138 (96.4%); Social interaction 135/138 (97.8%).

Factor structure

Factor analysis of the remaining 16 items produced three components with an eigenvalue <1.0 which explained 65.8% of the variance. The loadings on these three components are shown in Table 3. Component 1 (seven items) measures the effects of foot pain in relation to walking and standing (the 'Walking/standing domain'). Component 2 (five items) measures the level of foot pain (the 'Pain domain'). The two activities and two self-consciousness items loaded on Component 3 (the 'Social interaction domain').

Dimensionality

Application of a unidimensional Rasch model confirmed that the 16-item MOXFQ questionnaire was not unidimensional. However, when Rasch models were fitted to each of the three subscales (or domains), identified by the factor analysis, each was found to be unidimensional. (Further details of the Rasch analysis are provided in Appendix 2 at the end of this paper).

The 16-item MOXFQ (final version)

Scores for each domain are calculated as the sum of each individual item score. In each case, this is expressed on a metric of 0 – 100[1] (higher score representing greater severity)

Descriptive statistics from the three domains were: *Walking/standing domain* (seven items): mean 42.0 (SD 26.1), median 42.9, range 0–100; *Pain domain* (five items): mean 51.2 (SD 21.7), median 55.0, range 0–95; *Social interaction domain* (four items): mean 35.7 (SD 20.2), median 37.5, range 0–88.

$$\frac{100}{\text{Maximum possible domain scorce}} \times \text{Actual score}$$

Internal reliability (consistency)

The Cronbach's alpha coefficients for the Walking/standing and Pain domains were 0.92 and 0.86 respectively and the alphas for deletion of each item revealed that no one item had a particularly large effect on the reliability of each domain. Thus each factor had a high level of internal reliability [25, 26]. The reliability of the four-item Social Interaction domain was slightly less, at 0.73.

Convergent and divergent validity

Table 4 shows correlations between the MOXFQ domains, the AOFAS hallux clinical scale and SF-36 scores. All of these were negative (the MOXFQ scale is scored in an opposite direction from the other scales), demonstrating that the greater the MOXFQ scores (i.e. the worse the foot problems), the lower, (i.e. poorer) the AOFAS hallux clinical scale and SF-36 scores. The convergent validity of the MOXFQ was demonstrated by high correlations (i.e. $r > 0.5$, shown in bold typeface in Table 4) between the AOFAS hallux clinical scale and the MOXFQ Walking/standing and Pain domains; the correlation with the Social interaction domain was lower. The MOXFQ Walking/standing domain was strongly associated with the SF-36 Physical functioning, Role Physical, and Pain domains, and with the SF-36 PCS score. The MOXFQ Pain domain was most strongly related to the SF-36 Pain domain and to the PCS score.

Divergent validity of the MOXFQ was shown by the low correlations between the MOXFQ Walking/standing and Pain domains and the SF-36 Role emotional, Social functioning, Mental health, Energy/vitality and General health perception domains as well as the SF-36 MCS score. The Social interaction domain was not strongly related to any SF-36 domain, neither was it strongly related to the AOFAS hallux clinical scale score.

[1] Conversion to metric score:

Table 3. Factor analysis: rotated loadings of each of the 16 items on the three factors (items ordered by loading size) representing the final version of the MOXFQ

Original item no.	New item no.	MOXFQ item[a]	Factor 1: Walking/standing	Factor 2: Pain	Factor 3: Social interaction
4.	4.	Walk slowly	**0.866**	0.144	−.028
5.	5.	Stop and rest	**0.765**	0.421	0.047
2.	2.	Avoid walking long distances	**0.760**	0.306	0.137
6.	6.	Avoid some hard/rough surfaces	**0.732**	0.242	0.166
8.	8.	Catch bus/use car instead of walking	**0.717**	0.185	0.281
3.	3.	Change the way I walk	**0.678**	0.329	0.172
7.	7.	Avoid standing for a long period of time	**0.657**	0.484	0.219
16.	12.	Shooting pains	0.188	**0.759**	−0.039
15.	11.	Pain worse in the evening	0.208	**0.755**	0.155
20.	16.	Pain in bed at night	0.201	**0.750**	0.144
19.	15.	Usual level of pain	0.446	**0.742**	0.093
13.	1.	Pain in my foot	0.347	**0.739**	0.092
12.	10.	Self-conscious about shoes	0.040	−0.052	**0.816**
11.	9.	Self-conscious about feet	0.003	0.112	**0.803**
18.	14.	Cannot do all my social/recreational activities	0.412	0.211	**0.622**
17.	13.	Cannot carry out work/everyday activities	0.396	0.172	**0.562**
		% of variance explained	28.9%	22.6%	14.3%

[a]See Table 5 for precise wording of items.

Table 4. Spearman correlation coefficients between three MOXFQ domains and AOFAS hallux clinical scale, SF-36 domains, and SF-36 MCS and PCS scores

MOXFQ domains	AOFAS hallux scale	SF-36 domains								SF-36 components	
		Physical functioning	Role physical	Role emotional	Social functioning	Mental health	Energy/ vitality	Pain	General health	PCS	MCS
Walking/ standing	**−0.556*****	**−0.677*****	**−0.579*****	−0.347***	−0.451***	−0.096	−0.331***	**−0.543*****	−0.228**	**−0.634*****	−0.432***
Pain	**−0.597*****	−0.457***	−0.374***	−0.330***	−0.359***	−0.112	−0.276**	**−0.528*****	−0.125	**−0.521*****	−0.367***
Social interaction	−0.285**	−0.367***	−0.308***	−0.168	−0.134	−0.059	−0.047	−0.273**	−0.038	−0.219*	−0.113

High correlation values ($r > 0.5$) are shown in bold typeface.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Responsiveness

Table 5 shows the mean pre- and 12 month post-operative scores, change scores and effect sizes comparing MOXFQ, AOFAS hallux and SF-36 domains. Each of the assessments exhibited statistically significant mean changes in scores at 12 months following surgery. All significant changes were in the direction of improvement (reduced severity). Effect sizes (representing the magnitude of change) for each of the 3 MOXFQ scales and the AOFAS hallux clinical scale were > 1.0, indicating a very large degree of change following surgery. The generic scales of the SF-36 produced lower effect sizes than the foot-specific measures, indicating that the foot-specific scales were more responsive in this context. Only the SF-36 pain and physical function domains registered moderate to high effect sizes.

Discussion

There is an urgent need for scientific evaluation of foot and ankle surgery which in turn requires the employment of appropriate (patient-focused) standard methods of outcomes assessment [2, 3, 28, 30–32]. Outcome measures need to be validated

Table 5. Instrument responsiveness: mean pre- and 12 month post-operative scores, change scores and effect sizes comparing MOXFQ, AOFAS hallux and SF-36 domains

	Pre-surgery		12 Months post-surgery				Significance test[a]	Effect size[b]
	Mean	SD	Mean	SD	Mean change	SD	*p*	
MOX-FQ scales								
Foot pain (n = 117[c])	52.65	20.90	19.87	20.87	32.78	25.09	< 0.001	1.57
Walking/standing (n = 120)	44.85	25.37	16.37	23.07	28.48	27.58	< 0.001	1.12
Social interaction (n = 119)	46.59	22.97	11.82	18.73	34.77	24.06	< 0.001	1.51
AOFAS								
Hallux scale (n = 107)	54.80	11.92	84.15	15.03	−29.35	16.82	< 0.001	−2.46
SF-36								
Bodily pain (n = 119)	61.62	24.37	77.22	21.20	−15.59	26.57	< 0.001	−0.64
Physical function (n = 109)	75.00	22.64	85.28	18.67	−10.28	19.21	< 0.001	−0.45
Role-physical function (n = 120)	75.10	27.00	85.52	24.71	−10.42	31.02	< 0.001	−0.39
Mental health (n = 122)	70.61	16.99	77.50	16.28	−6.89	14.41	< 0.001	−0.41
Role-mental function (n = 120)	83.00	23.16	91.00	18.38	−7.10	25.48	0.001	−0.35
Vitality/energy (n = 121)	57.33	20.42	62.50	18.13	−5.17	17.89	0.002	−0.25
Social function (n = 116)	78.34	23.33	84.91	21.27	−6.57	25.40	0.006	−0.28
Health perception (n = 112)	76.18	19.41	79.64	17.23	−3.46	14.86	0.015	−0.18

[a]Paired *t*-tests.
[b]The positive/negative direction of the sign relates to the direction in which each scale is scored to denote increasing severity and is irrelevant to the size of the effect.
[c]The n represents the number of cases with complete data provided at both pre- and post-operative assessments to permit the evaluation of scales and thus change in pain, function, overall health-related quality of life.

in the context in which they are to be used. We concentrated on one context – surgery for hallux valgus – because hallux valgus is an extremely common condition [8, 9], (particularly in women), and the need for trials to evaluate treatment is particularly pressing in this area [2, 10].

The 16 item questionnaire developed in this study (the 'MOXFQ') represents the evolution of the MFPDQ – a questionnaire previously developed to assess foot pain and function [11] – into a new outcome measure for foot surgery. The new questionnaire addresses themes that were originally identified from interviews conducted with people who had various foot problems, in a largely community based study [11]. These same broad themes were confirmed as relevant in the context of surgery for hallux valgus in the current study. However, some amendments and an increased emphasis on the measurement of pain were found to be necessary.

The methods that we have used are well described and widely accepted [33, 34]. Formally tested on patients recruited to a prospective study, the new 16-item version of the questionnaire appeared very acceptable to patients. Factor analysis was used to select items for inclusion in each dimension of the questionnaire and Rasch analysis confirmed unidimensionality and that there was no redundancy in the item set.

Factor analysis revealed three subscales underlying the 16 items (a Foot Pain domain, a Walking/standing domain and a Social interaction domain), each of which is unidimensional and internally consistent (denoted by Cronbach's alpha).

The assessment of convergent and divergent validity (correlation with clinical foot assessment scores and SF-36 general health domains) revealed that the new questionnaire performed satisfactorily. The MOXFQ Walking/standing and Pain domains correlated highly with the AOFAS hallux clinical scale as well as with the relevant domains of the – generic – SF-36, as hypothesised. Correlations with the MOXFQ Social Interaction domain were lower – which was unsurprising as this scale places some emphasis on 'cosmesis' i.e. (in

this context) patients' feelings of self-consciousness about the appearance of their foot and the kind of footwear that they are restricted to wearing within social contexts. This is a particularly subjective perspective in relation to foot problems and one where the patients' perspective may well not mirror that of the clinician. In addition, the social function dimension of the SF-36 refers solely to social activities, and was not designed to measure any aspects of 'cosmesis'.

The MOXFQ has been demonstrated to be reliable, valid, and with a high rate of completion. Responsiveness (also referred to as sensitivity to change) is considered to be a particularly important property of a health outcome measure [24]. Tested within the context of patients undergoing surgery for hallux valgus – the MOXFQ was also demonstrated to be highly responsive. Further studies are now required to examine how well this questionnaire will perform in the context of surgery for other conditions affecting the foot and ankle.

Acknowledgements

We should like to thank The BUPA Foundation (the medical charity) for their generous support of this research project.We should also like to thank Adam Garrow for providing useful information and for his encouragement in the early days of this study. In addition, we very much appreciate being granted permission by Adam Garrow and collaborators who developed the earlier MFPDQ [11] to call the new measure the Manchester-Oxford Foot Questionnaire (MOXFQ).

Appendix 1.

The Manchester-Oxford Foot Questionnaire (MOXFQ)[2]

Table A1. Manchester-Oxford Foot Questionnaire (MOXFQ)

Circle as appropriate: **Right/Left Foot**[1] *Please tick √ one box for each statement*

During the past 4 weeks this has applied to me:	**None of the time**	**Rarely**	**Some of the time**	**Most of the time**	**All of the time**
1. I have pain in my foot	☐	☐	☐	☐	☐
2. I avoid walking long distances because of pain in my foot	☐	☐	☐	☐	☐
3. I change the way I walk due to pain in my foot	☐	☐	☐	☐	☐
4. I walk slowly because of pain in my foot	☐	☐	☐	☐	☐
5. I have to stop and rest my foot because of pain	☐	☐	☐	☐	☐
6. I avoid some hard or rough surfaces because of pain in my foot	☐	☐	☐	☐	☐
7. I avoid standing for a long time because of pain in my foot	☐	☐	☐	☐	☐
8. I catch the bus or use the car instead of walking, because of pain in my foot	☐	☐	☐	☐	☐
9. I feel self-conscious about my foot	☐	☐	☐	☐	☐
10. I feel self-conscious about the shoes I have to wear	☐	☐	☐	☐	☐
11. The pain in my foot is more painful in the evening	☐	☐	☐	☐	☐
12. I get shooting pains in my foot	☐	☐	☐	☐	☐
13. The pain in my foot prevents me from carrying out my work/everyday activities	☐	☐	☐	☐	☐
14. I am unable to do all my social or recreational activities because of pain in my foot	☐	☐	☐	☐	☐

15. During the past 4 weeks how would you describe the pain you usually have in your foot? *(please tick one box)*

None	Very mild	Mild	Moderate	Severe
❑	❑	❑	❑	❑

16. During the past 4 weeks have you been troubled by pain from your foot in bed at night? *(please tick one box)*

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
❑	❑	❑	❑	❑

[1]The foot to be assessed may be indicated here. Alternatively, each question may be customised to the right foot with all questions then repeated and customised to the left foot.

[2] Nb. The order, and hence the numbering, of items on the final 16-item version of the MOXFQ differs from all previous candidate versions described in this paper.

Appendix 2.

Dimensionality

Application of a unidimensional Rasch model to the 16 item questionnaire confirmed that the 16-item MOXFQ is not unidimensional.

Rasch models were next fitted to each of the three sub-scales, or domains, identified by the factor analysis. In this case, the seven items of the Walking/standing domain scale were clearly unidimensional ($\chi^2 = 12.04$, df = 14, $p = 0.603$), with no item significantly misfitting at the 5% level (highest residual = 1.6). In addition, the item thresholds were all properly ordered. The five-item pain domain was also unidimensional ($\chi^2 = 5.60$, df = 10, $p = 0.848$), with no item significantly misfitting at the 5% level (highest residual 0.9). The item thresholds, however, were not properly ordered for items 2, 17, and 18. The four-item 'Social interaction domain' was also found to be unidimensional ($\chi^2 = 3.86$, df = 8, $p = 0.869$), with no item significantly misfitting at the 5% level (highest residual = 0.5).

References

1. Mann RA, Plattner PF. Ankle and foot: Editorial overview. Curr Opin Orth 1990; 1: 111–112.
2. Ferrari J, Higgins JP, Prior TD. Interventions for treating hallux valgus (abductovalgus) and bunions. Cochrane Database Syst Rev 2004; 1: CD000964.
3. Parker J, Nester CJ, Long AF, Barrie J. The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery. Foot Ankle Int 2003; 24: 56–60.
4. Katsambas A, Abeck D, Haneke E, et al. The effects of foot disease on quality of life: Results of the Achilles Project. J Eur Acad Dermatol Venereol 2005; 19: 191–195.
5. Budiman ME, Conrad KJ, Roach KE. The Foot Function Index: A measure of foot pain and disability. J Clin Epidemiol 1991; 44: 561–570.
6. Bennett PJ, Patterson C, Wearing S, Baglioni T. Development and validation of a questionnaire designed to measure foot-health status. J Am Podiatr Med Assoc 1998; 88: 419–428.
7. AAOS Outcome Questionnaire. Musculoskeletal outcomes data evaluation and management system (MODEMS). Version 2000. 2000.
8. Wilson DW. Hallux valgus and rigidus. In: Helal B, Wilson D (eds.), The Foot. London: Churchill Livingstone, 1988: 411–483.
9. Dawson J, Thorogood M, Marks S-A, Juszczak E, Dodd C, Lavis G, et al. The prevalence of foot problems in older women: A cause for concern. J Public Health Med 2002; 24: 77–84.
10. Resch S, Stenstrom A, Jonsson K, Reynisson K. Results after chevron osteotomy and proximal osteotomy for hallux valgus: A prospective, randomised study. The Foot 1993; 3: 99–104.
11. Garrow AP, Papageorgiou AC, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. Development and validation of a questionnaire to assess disabling foot pain. Pain 2000; 85: 107–113.
12. Barrett P, Kline P. The observation to variable ratio in factor analysis. J Personal Group Behav 1981; 1: 23–33.
13. Ware-JE J, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473–483.
14. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. J Epidemiol Commun Health 1999; 53: 46–50.
15. Ware JE, Kosinski M, Keller S. SF-36 Physical and Mental Summary Scales: A User's Manual. The Health Institute, New England Medical Center: Boston, Massachusetts, 1994.
16. Kitaoka H, Alexander I, Adelaar R, Nunley J, Myerson M, Sanders M. Clinical rating systems for ankle-hindfoot, midfoot, hallux and lesser toes. Foot Ankle Int 1994; 15: 349–353.
17. Guyton GP. Theoretical limitations of the AOFAS scoring systems: An analysis using Monte Carlo modeling. Foot Ankle Int 2005; 22: 779–787.
18. SPSS Inc. SPSS Statistical Software. Release 11.5. Headquarters, 233 S. Wacker Drive, 11th floor Chicago, Illinois 60606, USA.: SPSS Inc., 2002.
19. Rasch Unidimensional Measurement Models. RUMM2010. 14 Dodonaea Court, Duncraig WA 6023, Australia., RUMM Laboratory Pty Ltd, 1998.
20. Prieto L, Alonso J, Lamarca R. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. Health Qual Life Outcomes 2003; 1: 1–27.
21. Wolfe F, Kong SX. Rasch analysis of the Western Ontario MacMaster Questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis 1999; 58: 563–568.
22. Hambleton RK, Jones RW. Comparison of classical test theory and item response theory and their applications to test development. Educ Measure Issues Practice 1993; 12: 38–47.
23. Andrich D. Rasch Models for Measurement. London: Sage Publications, 1988.
24. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993; 118: 622–629.
25. Nunnally JC, Bernstein IH. Psychometric Theory. New York: Columbus, OH: McGraw-Hill 1978.
26. Kline P. A Handbook of Psychological Testing. London: Routledge, 1993.
27. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27: S178–S189.
28. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Evaluation of the SF-36 Version II in the United Kingdom. J Epidemiol Commun Health 1999; 53: 46–50.
29. Royal College of Surgeons of England. Background to outcomes database. http://www.rcseng.ac.uk/surgical/research/ceu/projects_ongoing/proj_outcomes_html. 2004.

30. O'Doherty DP, Lowrie IG, Magnussen PA, Gregg PJ. The management of the painful first metatarsophalangeal joint in the older patient. Arthrodesis or Keller's arthroplasty? J Bone Joint Surg [Br] 1990; 72: 839–842.
31. Amadio PC. Outcomes measurements [editorial]. J Bone Joint Surg [Am] 1993; 75: 1583–1584.
32. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. Br Med J 1992; 305: 1074–1077.
33. Streiner DL, Norman GR. Health Measurement Scales – A Practical Guide to their Development and Use. New York: Oxford University Press, 1989.
34. Fayers PM, Machin D. Quality of Life – Assessment, Analysis and Interpretation. Chichester: John Wiley & Sons Ltd, 2000.

Address for correspondence: Dr Jill Dawson, Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford OX37LF, UK
Phone: +44-01865-227136; Fax: +44-01865-226711
E-mail: jill.dawson@dphpc.ox.ac.uk