ELSEVIER

Journal of Clinical Epidemiology 58 (2005) 47–55

Journal of Clinical Epidemiology

OAKHQOL: A new instrument to measure quality of life in knee and hip osteoarthritis

Anne-Christine Rat[a], Joël Coste[b], Jacques Pouchot[c], Michèle Baumann[a], Elizabeth Spitz[d], Nathalie Retel-Rude[e], Janine-Sophie Le Quintrec[f], Dominique Dumont-Fischer[g], Francis Guillemin[a,*]

[a]*EA 3444, Service d'Epidémiologie et Evaluations Cliniques, Hôpital Marin - CHU de Nancy, Nancy, France*
[b]*Département de Biostatistique et d'Informatique Médicale, Hôpital Cochin, Paris, France*
[c]*Service de Médecine Interne, Hôpital Louis Mourier, Colombes, France*
[d]*Laboratoire de Psychologie de la Santé, UFR Sciences Humaines et Arts, Université de Metz, France*
[e]*Laboratoire de Biostatistiques Pharmaceutiques, Faculté de Médecine et de Pharmacie, Université de Franche-Comté, Besançon, France*
[f]*Services de Rhumatologie et d'Orthopédie, Hôpital Cochin, Paris, France*
[g]*Service de Rhumatologie, Hôpital Avicenne, Bobigny, France*

Accepted 2 April 2004

Abstract

Objective: To develop a questionnaire with which to measure quality of life (QoL) in patients with knee and hip osteoarthritis (OA).

Study Design and Setting: Thirty-two caregivers and 96 OA patients were interviewed individually (using cognitive and face-to-face techniques) and in focus groups. A group of experts working independently at first and then consensually used the interview transcripts to generate a 46-item questionnaire.

Results: Analysis of questionnaires completed by 263 patients with hip or knee OA resulted in the exclusion of three items (two because of low reliability and one because of a low response rate). Principal component analysis revealed four factors: physical activity, mental health, social functioning, and social support. A pain dimension was individualized. Preliminary testing showed the reliability of the five dimensions to be satisfactory (intraclass correlation coefficients: 0.70–0.85), construct validity was adequate when correlated with the SF36 (Spearman correlation coefficients: 0.43–0.75), and discrimination was satisfactory. The osteoarthritis knee and hip quality of life questionnaire (OAKHQOL) consists of 43 items in five dimensions and three independent items.

Conclusion: The OAKHQOL is the first specific knee and hip OA quality of life instrument. Its development followed an *a priori* structured strategy to ensure content validity. It meets psychometric requirements for validity and reliability.

Keywords: Quality of life; Osteoarthritis; Hip; Knee; Psychometry

1. Introduction

Osteoarthritis (OA) is a chronic disease and a major cause of pain and disability. An estimated 7–11% of the populations of developed countries have symptomatic OA, and 27–44% have radiographic disease; the annual incidence in the United Kingdom is reported to be 3.1% [1–3]. As OA increases in prevalence with age, the growing proportion of elderly people in the populations of many countries will lead to it becoming an increasingly important global public health problem. OA already accounts for considerable expenditure by health care providers [4,5], and the cost to OA patients themselves can be substantial [6]. Apart from its economic consequences, OA is a major cause of disability, quality of life impairment, and social dysfunction [6]. It should therefore be a high priority for health care professionals, researchers, and public health decision-makers.

There is a growing interest in using quality of life (QoL) assessment to help investigate the impact of new pharmaceutical products and other interventions. The US FDA guide for clinical development programs in OA recommends that efficacy endpoints include: a measure of pain, patient global assessment, and a self-administered questionnaire covering pain and function (McMaster Western Ontario questionnaire [WOMAC] or Lequesne's index) [7,8]. QoL instruments are particularly valuable in patient global assessment because of their ability to capture more than just pain and disability [9].

* Corresponding author. Tel.: +33-38-385-2163; Fax: +33-38-385-1205.
E-mail address: Francis.Guillemin@sante-pub.u-nancy.fr (F. Guillemin).

0895-4356/04/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi: 10.1016/j.jclinepi.2004.04.011

Several tools have been used to measure health-related quality of life (HRQoL) among patients with OA of the lower limbs [7,10] and those undergoing total hip or knee surgery [11–16]. The Medical Outcomes Study Short-Form 36 (SF36) has been widely applied but, as a generic instrument, tends to be less responsive than specific instruments [17], particularly in the context of a medical or rehabilitation intervention rather than joint replacement. Comparisons of the SF36 with a disease-specific instrument (WOMAC) in patients undergoing knee replacement surgery report that they measure different aspects of health and should probably be used together [12,13]. However, no specific HRQoL instrument has been developed for patients with OA of the lower limb. WOMAC [7] and Lequesne's index [8] measure pain and functional disability but do not take other domains of QoL into account. The Arthritis Impact Measurement Scales (AIMS2) tool [18] and its short form AIMS2-SF [19] have been considered for use in OA but have a limited usefulness among patients with a high prevalence of lower limb disability [20].

There is a clear need for a disease-specific instrument with good content, construct validity, and responsiveness in assessing the QoL of patients with lower limb OA. We hypothesized that specific aspects of QoL are encountered by patients with knee and hip OA. Apart from affecting physical activities, knee or hip OA may also have an impact on mental health in terms of anxiety and depression, on sleep, on sexuality, and on social functioning [21,22]. Some specific aspects like social support have also already been demonstrated to be of importance in this pathology [23]. The combination of the SF36 with the WOMAC or the Lequesne index may thus not capture these specific aspects of HRQoL expressed by patients with knee and hip OA. The knee and hip osteoarthritis quality of life questionnaire (OAKHQOL) should be more able to apprehend aspects specifically appropriate to knee and hip OA patients.

Our aim was to develop a new instrument (OAKHQOL) with the ability to capture a patient's perception of his or her disease and with the psychometric properties required for use in clinical trials and observational studies. As the construction process of a QoL instrument has a major influence on its content and construct validity, OAKHQOL was developed *a priori*. The present article reports its conception, development, and early testing.

2. Materials and methods

OAKHQOL was developed in three stages: first, a qualitative stage to define the concept and content of the instrument, elicit verbatim remarks, build categories, and generate items; second, a quantitative stage to examine the properties of items, and determine the dimensional and factorial structure of the finished questionnaire; and third, preliminary psychometric analysis (item-scale correlations, construct validity, reliability).

2.1. Qualitative stage

2.1.1. First step: to define the concept of OAKHQOL

This step involved experts in rheumatology, psychology, sociology, and QoL, and patients with lower limb OA. The overall concept and its components were based on the World Health Organization (WHO) definition of health and quality of life [24]. The International Classification of Functioning disability and health (ICF) framework was used to categorize different aspects of health [25,26].

2.1.2. Second step: to elicit relevant verbal material

Patients with hip or knee OA, and relevant health professionals, were recruited to take part in focus groups and individual interviews. The sample of patients was drawn from those attending rheumatology and orthopedic surgery outpatient clinics and balanced using a quota method with stratification for age, sex, OA location (hip or knee), and medical or surgical stage of OA. This sampling method was not representative of the epidemiology of the disease but was chosen for its ability to elicit comprehensive and relevant items. Health professionals were drawn from disciplines familiar with the various stages of hip and knee OA and the options for management. They included rheumatologists, orthopedic surgeons, rehabilitation specialists, physiotherapists, and occupational therapists.

Verbal expressions used by patients and health professionals in this context were elicited in five ways [27]. The first was to interview patients using a cognitive technique based on memory retrieval, knowledge representation, and communication [28]. The second source was spontaneous utterance in traditional face-to-face interviews with patients. Sources three and four were focus groups of patients and health professionals, and five was face-to-face interviews with health professionals involved in the care of OA patients.

2.1.3. Third step: content analysis and generation of items

Six health sociologists and psychologists working independently in pairs conducted a semantic theme content analysis of the tape-recorded and transcribed interviews based on the pre-specified conceptual framework. Verbatim remarks extracted from the transcripts were grouped into categories agreed by consensus between pairs. Items were formulated accordingly, and any that were duplicated or irrelevant to QoL were discarded (e.g., coping strategies, satisfaction with care, and drug dependency were not retained). Verbatim and derived items were presented to a panel of experts (health psychologists and sociologists, rheumatologists, rehabilitation specialists, orthopedists, linguists, methodologists, epidemiologists, OA patient self-help group members, and a consumer representative) who selected a manageable number. The choice was guided mainly by content, but also took account of the frequency with which items were mentioned in interviews, their importance to patients, and their relevance to the concept of QoL.

Response modalities, how items were expressed, and the wording of instructions to those completing the questionnaire were also determined by consensus among the experts. Respondents to OAKHQOL are instructed to consider how OA has affected their QoL during the previous 4 weeks. The response format and timeframe of reference is consistent throughout in order to maximize precision and increase acceptability. Each item is measured on a numerical rating scale from 0 to 10 and the mean item score becomes the corresponding dimension score.

2.2. Quantitative stage

The object of this stage was to document psychometric properties of relevance in selecting items for the final version of OAKHQOL.

Patients were recruited in six outpatient clinics by rheumatologists and orthopedic surgeons from two distinctive areas of France (the Lorraine region and Paris) in an attempt to represent a wide spectrum of clinical features and ways of life. Subjects were required to have OA according to American College of Rheumatology criteria [29,30], to speak French, and to be free of any other disabling disorder. Three groups were defined according to the severity of their disease: group 1, patients managed medically; group 2, those scheduled to receive prosthetic replacement surgery within 3 months; and group 3, patients who had undergone hip or knee arthroplasty within the previous 2 years.

All participants provided written informed consent. They were asked to fill in the OAKHQOL 1.0 and the SF36 questionnaires, and provided socio-demographic and clinical data. A second questionnaire was mailed 10 days later to assess reproducibility.

Chi-square, Fisher's exact test, and ANOVA were used to test for differences in socio-demographic and clinical characteristics between the three groups of patients. Scores for each of the eight dimensions of the SF36 were calculated according to standard procedures [31]. The percentage of missing data was tabulated for each item. Less than 5% was considered acceptable [32,33]. A higher proportion of missing data was accepted for items considered more intrusive (those related to sexual activity, for example).

The frequency distributions of individual items were examined to determine if all the response modalities were used and to detect ceiling or floor effects. There is no rule concerning the acceptable maximal values of extreme modalities because the magnitude of ceiling or floor effects depends on the number of response modalities. A uniform frequency distribution reflects substantial variability in response choices—a desirable characteristic to be expected of a sample incorporating very different levels of severity of OA. A high percentage of extreme responses is acceptable for items that explore a high or a low level of difficulty. For instance, a high percentage of respondents to the item "I need a stick to walk" would be expected to answer 0 ("never"). Nevertheless, it is still important to include this type of item in the scale in order to measure the full range of the construct.

Items were assessed for reliability in terms of intraclass correlation coefficients (ICCs) [34] derived from a two-way analysis of variance in a random effect model, and by the Bland and Altman graphical method [35]. Patients with significant changes to their clinical status between the two administrations were excluded from the reproducibility analysis. An ICC of more than 0.80 indicated excellent reproducibility, one between 0.61 and 0.80 moderate reproducibility, and one between 0.41 and 0.60 fair reproducibility [34].

Dimensionality and factorial structure were initially investigated using principal component analyses (PCA). The number of factors to be retained was determined based on screeplot and eigenvalues $\geqslant 1$ and orthogonal and oblique promax rotations were then performed. These analyses were exploratory in order to reveal whether items supposed to belong to the same scale grouped coherently. Three items excluded in the PCA were retained as separate questions. They were expected to concern only a minority of respondents and would have jeopardized the validity of the PCA since too many observations would have been removed. The items were: "My professional activity is affected," "My relationship with my partner is affected," and "My sexual activity is limited."

Statistical analyses were performed using Statistical Analysis System version 8 for Windows (SAS institute, Cary, NC).

2.3. Preliminary psychometric analyses of the instrument: item-scale correlation analyses, construct validity, and reliability

Scores for each scale were obtained by calculating the means of the corresponding items. Scale scores were calculated only if more than half of the items were completed, and normalized to range from 0 (best QoL) to 10 (worst QoL).

Analyses based on the multitrait analysis program revised (MAP-R) were conducted according to the program developed for The Medical Outcomes Study [36,37]. The MAP-R allows for the calculation of item-scale correlations (corrected for overlap) and correlations among scales. Item-scale correlation coefficients of 0.40 or more are acceptable.

Cronbach alpha coefficients were calculated in order to assess scale consistency, and correlations between scales were used to evaluate the degree of distinction between them.

Construct validity was investigated using Spearman correlation coefficients for convergent and divergent validity and non-parametric tests (Kruskal Wallis) for discriminant validity (the parameters studied did not show a normal distribution). Construct validity was assessed by correlating OAKHQOL scale scores with the corresponding SF36 scales and with the pain visual analog scale (VAS) and by comparisons of groups differing by age, sex, body mass index (BMI), and OA severity. It was hypothesized that older patients, females, patients with a higher BMI, and patients at a surgical

stage of severity would have higher OAKHQOL scale scores (worse QoL). Scales were assessed for reliability in terms of ICC.

3. Results

3.1. Qualitative stage

One hundred twenty-eight patients and 32 health professionals participated in this stage. The initial analysis of interview transcripts identified 119 relevant items, of which, 83 were pertinent and related to QoL. The first version of the OAKHQOL 1.0 was a self-administered questionnaire of 46 items reported in terms of their frequency and intensity.

3.2. Quantitative stage

3.2.1. Sample characteristics

Of 263 patients recruited in outpatient clinics, 139 were being treated medically, 97 were scheduled for surgery, and 27 had undergone total arthroplasty of hip or knee within the previous 2 years. Among them, 125 had participated in the qualitative stage of OAKHQOL's development.

Socio-demographic and clinical characteristics were as shown in Table 1. The sample was predominantly female (59%) with a mean age of 66 years. Eighty percent of patients were not working, 73% were married, and 46% lived in rural areas. The primary site of lower limb OA was the hip in 44% and the knee in 56%. Differences in clinical variables between the surgical stage group and the other groups were as expected and provided the heterogeneity required to test the instrument.

3.2.2. Descriptive statistics of items

Proportions of missing data are shown in Table 2. Only one item was missing more than 5% of data (Q17). Four items allowed for a response of "not applicable." They were pertinent to 26–73% of patients and the corresponding proportion of missing data was greater than 5%. The items concerned related to use of public transport, employment, relationships, and sexual activity. However, the proportions of missing data were not as high as might be expected.

Responses were evenly distributed along the scale for most items. Thirteen items had an extreme response that was chosen by more than 30% of respondents (Table 2). Mean scores revealed that the disease had a considerable impact on the health status of respondents as expressed in most items and the scales identified later (see below).

Test-retest reliability data are shown in Table 2. Of the 263 patients who completed the initial questionnaire, 77% (203) returned the second one mailed 10 days later. Only questionnaires returned between 10 and 21 days after the first was completed were retained, leaving 161 pairs for analysis of reproducibility. Six items had an ICC of less

Table 1
Socio-demographic and clinical characteristics of OA patients comparing three groups of severity

		Medical stage n = 139	THA or TKA[a] n = 27	Surgical stage n = 97	p
Sex	Female	92 (67.1)	16 (59.2)	47 (48.4)	0.02
Age years (mean ± SD[b])		65.1 ± 10.6	64.0 ± 10.4	68.6 ± 8.9	0.002
Marital status	Living as a couple	86 (67.7)	20 (80.0)	65 (79.3)	0.13
Residence	Urban	93 (73.8)	17 (73.9)	15 (17.9)	<0.0001
Education	Primary	57 (42.9)	7 (25.9)	68 (82.9)	<0.0001
	Secondary	51 (38.3)	13 (48.1)	13 (15.8)	
	University	25 (18.8)	7 (25.9)	1 (1.2)	
Employment status	Employed	33 (25.2)	5 (20.8)	9 (11.1)	0.04
BMI[c] (kg/m2)	<25	66 (40.0)	63 (57.3)	49 (49.5)	0.03
	25–30	13 (7.9)	10 (9.1)	5 (5.0)	
	>30	86 (52.1)	37 (33.6)	45 (45.4)	
OA joint	Hip	33 (28.0)	15 (62.5)	58 (60.0)	<0.0001
	Knee	86 (72.0)	9 (38.5)	39 (40.0)	
Pain intensity (mean ± SD)	VAS[d]: 0–100	49.9 ± 26.4	—	76.5 ± 99.1	0.002
SF36[e] (mean ± SD)	Physical functioning	48.8 ± 23.5	56.6 ± 24.0	35.8 ± 21.4	<0.0001
	Role physical	41.2 ± 40.1	50.9 ± 41.9	22.0 ± 31.9	<0.0001
	Bodily pain	42.3 ± 20.8	54.1 ± 25.2	34.0 ± 16.7	<0.0001
	Mental health	59.0 ± 21.1	58.9 ± 21.3	52.3 ± 19.0	0.04
	Role emotional	49.6 ± 43.7	58.0 ± 40.9	26.9 ± 38.6	<0.0001
	Social functioning	64.9 ± 25.1	61.1 ± 21.5	62.2 ± 24.6	0.06
	Vitality	43.8 ± 19.8	48.4 ± 18.0	37.1 ± 15.0	0.002
	General health	51.2 ± 19.8	57.6 ± 17.3	54.3 ± 16.5	0.15

Numbers and percentages unless indicated.

[a] THA or TKA: patients with total hip arthroplasty (THA) or total knee arthroplasty (TKA) within the previous 2 years.

[b] SD: standard deviation.

[c] BMI: body mass index (kg/m^2).

[d] VAS: visual analog scale.

[e] SF36: scores in the range 0 to 100 with a lower score indicating greater distress.

Table 2
Floor, ceiling effects, missing, "not concerned" data, and test-retest intraclass correlation coefficients of the items of the OAKHQOL 1.0

	Abbreviated item content of OAKHQOL 1.0	Floor effect (%)	Ceiling effect (%)	Not concerned (%) n	Missing data (%) n	ICC[a]	95% confidence interval
Q1	Walking	4.9	9.5	—	(0.0) 0	0.63	0.53–0.72
Q2	Bending or straightening	3.6	13.6	—	(0.7) 2	0.74	0.66–0.81
Q3	Carrying heavy things	4.6	12.8	—	(1.9) 5	0.68	0.58–0.75
Q4	Going down stairs	9.5	9.9	—	(0.4) 1	0.76	0.68–0.81
Q5	Climbing stairs	4.5	11.3	—	(0.4) 1	0.75	0.67–0.81
Q6	Taking a bath	18.7	12.1	—	(3.7) 10	0.78	0.71–0.83
Q7	Dressing	14.5	9.9	—	(0.4) 1	0.76	0.68–0.82
Q8	Cutting toe-nails	9.1	21.4	—	(0.7) 2	0.76	0.68–0.82
Q9	Staying for a long time in the same position	4.5	12.7	—	(1.1) 3	0.78	0.72–0.84
Q10	Getting moving after staying in the same position	2.7	12.2	—	(0.4) 1	0.71	0.62–0.78
Q11	Need a stick to walk	59.7	8.1	—	(0.4) 1	0.88	0.84–0.91
Q12[b]	Able to do the things one used to	8.6	6.8	—	(0.7) 2	0.36	0.21–0.49
Q13	Need help	44.4	6.5	—	(2.6) 7	0.81	0.75–0.86
Q14	Getting in and out a car	10.9	9.5	—	(0.4) 1	0.71	0.63–0.78
Q15	Using public transport	29.2	9.5	(34.0) 91	(6.2) 11	0.78	0.68–0.86
Q16[b]	Had usual physical leisure activities	23.3	13.5	—	(4.1) 11	0.45	0.31–0.57
Q17[b]	Had usual social and leisure activities	32.7	15.6	—	(6.7) 18	0.59	0.47–0.69
Q18[c]	Hindered in professional activity	25.5	9.8	(74.0) 199	(14.7) 10	0.71	0.45–0.86
Q19	Need to spare oneself	6.0	7.8	—	(3.0) 8	0.72	0.64–0.79
Q20	Take longer doing things	5.5	8.7	—	(1.9) 5	0.81	0.74–0.86
Q21	Intensity of pain	3.2	5.9	—	(2.2) 6	0.65	0.55–0.73
Q22	Frequency of pain	1.8	12.4	—	(2.6) 7	0.69	0.59–0.76
Q23	Having difficulties getting to sleep because of pain	21.5	8.7	—	(1.1) 3	0.78	0.71–0.83
Q24	Wake up at night because of pain	22.4	9.1	—	(1.9) 5	0.76	0.68–0.82
Q25	Feel depressed because of pain	26.3	5.5	—	(2.2) 6	0.69	0.60–0.76
Q26	Feel older than my years	26.8	11.4	—	(1.5) 4	0.67	0.57–0.75
Q27	Been afraid of being dependent on others	15.4	23.2	—	(1.1) 3	0.76	0.69–0.82
Q28	Been afraid of becoming an invalid	14.5	25.4	—	(0.7) 2	0.73	0.65–0.80
Q29	Embarrassed when people see me	58.9	3.6	—	(1.9) 5	0.67	0.57–0.75
Q30	Worry	24.3	8.7	—	(2.6) 7	0.70	0.61–0.77
Q31	Feel depressed	37.6	4.1	—	(3.0) 8	0.68	0.58–0.75
Q32	Able to plan for the future	15.1	16.4	—	(1.5) 4	0.63	0.53–0.72
Q33	Wonder what is going to happen	26.4	9.1	—	(0.7) 2	0.78	0.72–0.84
Q34	Worried about the side-effects of treatment	28.0	8.3	—	(1.9) 5	0.68	0.58–0.75
Q35	Going out whenever would like	7.2	36.5	—	(0.7) 2	0.63	0.53–0.72
Q36	Have friends in whenever would like	5.4	36.8	—	(2.2) 6	0.64	0.53–0.72
Q37	Hindered in family life	36.0	4.2	—	(4.1) 11	0.59	0.48–0.69
Q38[c]	Hindered in life with partner	42.9	8.0	(27.2) 74	(5.2) 10	0.66	0.53–0.76
Q39[c]	Restricted in sexual life	38.7	6.7	(39.1) 104	(14.8) 24	0.71	0.57–0.81
Q40	Feel aggressive and irritable	23.2	1.4	—	(1.1) 3	0.74	0.66–0.80
Q41	Feel being a burden to close relatives	29.6	2.3	—	(3.7) 10	0.71	0.62–0.78
Q42	Talking about arthritis problems	0.9	36.7	—	(2.2) 6	0.58	0.46–0.67
Q43	Feel others understand arthritis problems	3.7	31.3	—	(2.6) 7	0.60	0.49–0.69
Q44	Feel embarrassed to ask for help	25.7	20.1	—	(4.1) 11	0.63	0.53–0.72
Q45	Feel support from people close to me	3.6	49.8	—	(2.2) 6	0.57	0.45–0.67
Q46	Feel support from people around	6.9	35.9	—	(3.0) 8	0.63	0.53–0.72
Dimensions of the OAKHQOL 2.0							
	Physical activities					0.84	0.79–0.88
	Mental health					0.85	0.80–0.89
	Pain					0.76	0.69–0.82
	Social support					0.70	0.61–0.77
	Social functioning					0.70	0.61–0.77

Abbreviated item content of the OAKHQOL are not the whole items; the translation and adaptation in English is still in process.

[a] ICC: intraclass correlation coefficient.

[b] Items that were deleted from OAKHQOL 1.0.

[c] Items not included in the principal component analysis and not included in the five dimensions of the OAKHQOL 2.0 (40 items in five dimensions plus three isolated items).

than 0.6 (values were close to 0.6 in four) and one of less than 0.4. Reliability errors of three of these items (Q12, Q16, and Q17) were non-homogeneous along the 0–10 scale range on the Bland and Altman graphic analysis.

3.2.3. Dimensionality and factorial structure

A four-factor PCA solution (Table 3) was retained based on screeplot and eigenvalues. Factor loading of one item (Q17) was very low on any of the four factors. Three items (Q12, Q16, Q17) were not considered further in the subsequent principal component analyses because of low reliability in two (Q12, Q16) and low response rate and absence of loading on any factor in one (Q17). The structure of the analysis obtained was identical with or without them. The first four factors explained 64% of the total variance.

The four scales were physical activities (19 items), mental health (14 items), social support (four items), and social functioning (three items). All items concerning pain loaded similarly on the physical factor and the mental health factor and were correlated to each other. The panel of experts therefore decided to include pain as an individual dimension. A PCA restricted to the physical activity and pain items clearly individualized all the latter in the same factor. The same result was obtained if pain items were combined with the mental health items in another separate analysis. In PCA with oblique rotation, the first two factors (physical activities and mental health) were indistinguishable, but the third and the fourth were independent.

Based on expert consensus, 43 items were retained in OAKHQOL 2.0, which comprises five dimensions, and three independent items concerning relationships, sexual activity, and professional life. The five scales and the three independent items will be used as separate outcomes.

3.3. Item-scale correlation analyses, construct validity, and reliability

The results of the MAP-R showed that correlation coefficients between items and their hypothesized scale were all greater than 0.4, other than item 32 (coefficient correlation: 0.39). Three items among the 43 loaded higher on their own scale but also on another subscale.

Cronbach alpha coefficients for the five scales were good (Table 4). The dimensions "social support" and "social functioning" were clearly distinct from the others. In contrast, as shown in the PCA, the pain scale was highly correlated to the physical and mental health subscales.

Table 4 shows the convergent and divergent validity of the OAKHQOL. The correlations were good or fair with the corresponding scales of the SF36 except for the social dimension, with which correlation was weak. VAS pain was weakly correlated with the OAKHQOL pain scale. The social support scale specific to OAKHQOL was weakly correlated with SF36 dimensions and VAS pain.

As expected, the physical activity scale scores showed statistically significant differences between the BMI classes and between the stages of OA. Discrimination was also good for the pain scale between BMI classes, sexes, and stages of OA. There were also differences between sexes for the mental health scale and between sexes and stages of OA for social functioning. No difference was seen between hip and knee OA (Table 5).

The test-retest correlation coefficients of the scales are shown in Table 2. Two dimensions (physical activity and mental health scales) had excellent reliability, and three moderate reproducibility.

Table 3
Distribution of items of OAKHQOL questionnaire according to PCA with orthogonal varimax rotation

	Abbreviated item content of OAKHQOL 1.0	Factor 1 Physical activities	Factor 2 Mental health	Factor 3 Social support	Factor 4 Social functioning
Q1	Walking	**0.82**			
…	…	…			
Q20	Take longer doing things	**0.69**			
Q21	Intensity of pain	**0.69**	0.49		
Q22	Frequency of pain	**0.59**	0.50		
Q23	Having difficulties getting to sleep because of pain	**0.60**	0.57		
Q24	Wake up at night because of pain	0.57	**0.59**		
Q25	Feel depressed because of pain	0.41	**0.72**		
Q26	Feel older than my years		**0.68**		
…	…		…		
Q44	Feel embarrassed to ask for help		**0.58**		
Q42	Talking about arthritis problems			**0.80**	
…	…			…	
Q46	Feel support from people around			**0.79**	
Q32	Able to plan for the future				**0.59**
…	…				…
Q36	Have friends in whenever would like				**0.63**

Significant loading of more than 0.4 are shown.
Sixty-four percent of the total variance was explained by the first four factors.
Items related to pain (Q21–Q24 loaded on both factor 1 and 2).
Abbreviated item content of the OAKHQOL are not the whole items, the translation and adaptation in English is still in process.

Table 4
Construct validity and correlations between the scales of the OAKHQOL and the scales of the SF36

	OAKHQOL scales				
	Physical activities	Mental health	Pain	Social Support	Social functioning
OAKHQOL (n = 263)					
Physical activities	(0.96)	0.57	0.70	0.07	0.46
Mental health		(0.92)	0.64	0.29	0.48
Pain			(0.91)	0.14	0.38
Social support				(0.81)	0.33
Social functioning					(0.73)
SF36 (n = 263)					
Physical functioning	**0.63**	0.46	0.48	0.17	0.52
Physical role	**0.50**	0.41	0.46	0.13	0.38
Bodily pain	0.65	0.53	**0.66**	0.22	0.43
Mental health	0.49	**0.75**	0.52	0.29	0.45
Role emotional	0.44	0.45	0.46	0.17	0.31
Social functioning	0.49	0.60	0.46	**0.34**	**0.43**
Vitality	0.58	**0.66**	0.59	0.17	0.44
General health	0.33	0.54	0.39	0.26	0.39
Pain VAS (n = 123)	0.53	0.34	**0.42**	0.09	0.35

Numbers in parentheses are Cronbach alpha coefficients.
Bold correlation coefficients indicate scales conceptually closed but not necessarily statistically significant.

4. Discussion

The OAKHQOL is the first specific knee and hip OA quality of life instrument. Its development followed an *a priori* structured strategy to ensure content validity. It meets psychometric requirements for validity and reliability.

The intention when developing OAKHQOL was to introduce elements of QoL that patients with hip and knee OA at various stages of severity report as affecting their daily lives. Within the framework provided by WHO, the *a priori* strategy adopted has allowed for the generation of a number of items unique to OA and its consequences.

Recommendations for the conduct of clinical drug trials highlight the importance of assessing the effects of intervention on non-signal (contralateral) joints and of patient global assessment. QoL instruments are therefore appropriate because of their ability to capture global effects [9].

Existing instruments for use in OA focus on symptoms and functioning, but provide no information about the overall perceived impact of the disease; patient-reported outcome instruments fail in that respect because they measure only functional status or are not specific to lower limb OA [7,8,13].

Comparison of a generic (SF36) and a disease-specific (WOMAC) instrument in the context of knee replacement surgery shows that they do not measure the same aspect of health and should probably be used together [12,13].

The concept of the OAKHQOL is based on the WHO definition of QoL. Content analysis of the tape-recorded transcripts elucidated different facets of the functionalist approach to QoL and of the needs-based model [38–42] and confirmed that patients expressed the impact of disease in terms of the needs they were unable to meet.

When comparing the OAKHQOL to the SF36, many themes were exclusive to the OAKHQOL (e.g.: social support, sleep, side effects of drugs, plan for the future, the embarrassment to be seen by people, the use of public transportation, the difficulty to move after staying in the same position, and sexuality). In this regard, 26 items were OAKHQOL exclusive. Among the 20 physical activity and pain items of the OAKHQOL, only 45% and 50% were part of the WOMAC and of the Lequesne index, respectively, and 4 of them, raised by more than 50% of patients during individual interviews, were completely new. Thus, the combination of WOMAC or Lequesne scales with the SF36 does not capture specific HRQoL of knee or hip OA as expressed by patients. The OAKHQOL is especially adapted to measure alterations of QoL specifically due to knee and hip OA.

Table 5
Construct validity of the OAKHQOL

		OAKHQOL scales															
			Physical activities			Mental health			Pain			Social support			Social functioning		
		N	Mean	(SD)	p	Mean	(SD)	p	Mean	(SD)	p	Mean	(SD)	p	Mean	(SD)	p
BMI	<25	89	5.0	(2.5)	0.0005	3.9	(2.4)	0.07	4.9	(2.6)	0.01	3.1	(2.4)	0.21	3.8	(2.8)	0.59
	25–30	92	4.7	(2.4)		3.4	(2.4)		4.7	(2.6)		2.5	(2.3)		3.6	(2.7)	
	>30	82	6.1	(2.1)		4.3	(2.4)		5.8	(2.6)		2.7	(2.5)		3.9	(2.4)	
Age	<60 years	74	5.1	(2.4)	0.36	4.5	(2.7)	0.07	5.7	(2.8)	0.07	2.9	(2.5)	0.74	3.5	(2.8)	0.49
	60–70 years	86	5.1	(2.3)		3.7	(2.4)		5.0	(2.6)		2.8	(2.5)		3.7	(2.8)	
	>70 years	103	5.5	(2.5)		3.6	(2.2)		4.8	(2.5)		2.6	(2.2)		3.9	(2.4)	
Sex	Female	155	5.5	(2.3)	0.06	4.2	(2.4)	0.02	5.5	(2.6)	0.005	3.0	(2.4)	0.09	4.1	(2.7)	0.008
	Male	106	4.9	(2.5)		3.5	(2.4)		4.6	(2.6)		2.5	(2.4)		3.2	(2.5)	
Joint	Hip	106	5.5	(2.6)	0.22	3.9	(2.5)	0.69	5.4	(3.0)	0.10	2.6	(2.4)	0.38	3.8	(2.8)	0.92
	Knee	134	5.1	(2.3)		3.8	(2.4)		4.8	(2.4)		2.7	(2.3)		3.6	(2.5)	
Severity	Post-surgery	27	3.7	(2.5)	<.0001	3.4	(2.7)	0.22	4.1	(3.0)	0.0006	1.8	(1.8)	0.008	2.6	(2.4)	0.04
	Medical stage	139	4.8	(2.4)		3.7	(2.4)		4.8	(2.6)		3.2	(2.5)		3.7	(2.8)	
	Surgical stage	97	6.4	(1.9)		4.2	(2.4)		5.9	(2.5)		2.4	(2.3)		4.0	(2.5)	

The scores range from 0 (best) to 10 (worst quality of life).

The OAKHQOL was developed in accord with the idea that the more similar data on the same concept are obtained using different approaches, the more rich and valid the findings [43,44]. The large sample of patients and experts who helped generate items, the variation in interview or focus group technique [27], and the analysis of transcripts from patients by six different analysts provide independent assessments of the same phenomena and minimize biases and measurement errors.

The carefully structured qualitative stage resulted in the deletion of only three items because of their metric properties. The approach to item selection adopted here is in agreement with the method developed by the French Quality of Life in Rheumatology group, which combines psychometric and clinimetric information, and creates a hierarchy of priorities [19,45] that favors the content of the item over its psychometric properties. The expert panel decided to retain the four items that concerned few patients because they reflected real life and were of value in exploring the broad spectrum of QoL impairment experienced by people with OA. Sexual functioning correlates relatively weakly with SF36 and is a good candidate for inclusion in specific instruments [31,46].

The OAKHQOL includes a pain dimension shown to load almost equally on the factors representing mental and physical health and not systematically on the physical component of SF36 [47]. Correlation between pain and function has also been shown to be weak in several other musculoskeletal diseases, including rheumatoid arthritis, hand OA, and low back pain [48,49].

This study has some limitations. First, the OAKHQOL so far has been developed and tested only in France and therefore its conceptual relevance and psychometric properties in other countries and cultures remains unknown. Cross-cultural adaptations according to published guidelines [50,51] are presently in development. Second, individual (and unavoidable) variability in the interpretation of interview transcripts cannot be completely excluded but is greatly limited here by assessment of the same phenomena by different teams. Third, longitudinal data are required to document sensitivity to change, and the OAKHQOL should also be tested in other samples.

Construct validity should be further investigated using other scales and clinical characteristics and a Rasch analysis is also underway to further investigate measurement properties. A study to assess validity, reliability, and responsiveness of the definitive version of the OAKHQOL on a new sample is in process. The OAKHQOL will be available for use in cohort studies measuring QoL in patients with lower limb OA, and trials of medical or rehabilitative interventions. It can also be used to assess preoperative predictors of QoL after hip or knee total replacement, and could aid in decision-making. An international group has been set up and is currently working to provide validated versions of the OAKHQOL in various languages including English.

Acknowledgments

The authors wish to thank M. Alfonsi, F. Arnould, X. Ayral, P. Bouillot, I. Cahitte, I. Chary-Valckenaere, P. Chaspoux, C. Cuny, J.P. Delagoutte, P. Fener, H. Gaudin, M.F. Gérard, F. Goupy Le Maitre, J. P. Gros, D. Hannouche, E. Hiraux, N. Jeanson, J.M. Lardry, D. Loeuille, D. Mainard, D. Nebout, G. Osnowycz, H. Ouakil, S. Perrot, S. Poiraudeau, J. Pourel, P. Prost, P. Rabany, M. Rousseau, Dr. Thomas, F Touzard, G. Vançon, L. Vastel, and J.P. Voilquin. This study was conducted with the support of the Clinical Epidemiology Center-Inserm (Institut national en santé et recherche medicale) of Nancy, the Ministry of Health, and the Nancy University Hospital. OAKHQOL 2.1 is available upon request.

References

[1] Felson DT, Zhang Y, Hannan MT, Naimark A, Weissman BN, Aliabadi P, Levy D. The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. Arthritis Rheum 1995;38:1500–5.

[2] Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41: 1343–55.

[3] Hart DJ, Doyle DV, Spector TD. Incidence and risk factors for radiographic knee osteoarthritis in middle-aged women: the Chingford Study. Arthritis Rheum 1999;42:17–24.

[4] Gabriel SE, Crowson CS, Campion ME, O'Fallon WM. Direct medical costs unique to people with arthritis. J Rheumatol 1997;24:719–25.

[5] Gabriel SE, Crowson CS, Campion ME, O'Fallon WM. Indirect and nonmedical costs among people with rheumatoid arthritis and osteoarthritis compared with nonarthritic controls. J Rheumatol 1997; 24:43–8.

[6] Lapsley HM, March LM, Tribe KL, Cross MJ, Brooks PM. Living with osteoarthritis: patient expenditures, health status, and social impact. Arthritis Rheum 2001;45:301–6.

[7] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.

[8] Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation–value in comparison with other assessment tests. Scand J Rheumatol Suppl 1987;65:85–9.

[9] FDA. Guidance for industry. Clinical development programs for drugs, devices and biological products intended for the treatment of osteoarthritis (OA) 1999. Available at: http://www.fda.gov/cber/gdlns/osteo. Accessed July 2, 2003.

[10] Pouchot J, Joubert J, Bono I. Apport des instruments de mesure d'incapacité fonctionnelle et de qualité de vie dans l'arthrose. [Contribution of functional disability and quality of life measures to the evaluation of patients with osteoarthritis.] Rev Rhum 1995;62:56S–63S.

[11] Mainard D, Guillemin F, Cuny C, Mejat-Adler E, Galois L, Delagoutte J. Evaluation à un an de la qualité de vie après arthroplastie totale de hanche et de genou. [Quality of life assessment one year after total hip or knee arthroplasty.] Rev Chir Orthop Reparatrice Appar Mot 2000;86:464–73.

[12] Bombardier C, Melfi CA, Paul J, Green R, Hawker G, Wright J, Coyte P. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. Med Care 1995;33: AS131–44.

[13] Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22:1193–6.

[14] Mangione CM, Goldman L, Orav EJ, Marcantonio ER, Pedan A, Ludwig LE, Donaldson MC, Sugarbaker DJ, Poss R, Lee TH. Health-related quality of life after elective surgery: measurement of longitudinal changes. J Gen Intern Med 1997;12:686–97.

[15] Lieberman JR, Dorey F, Shekelle P, Schumacher L, Kilgus DJ, Thomas BJ, Finerman GA. Outcome after total hip arthroplasty. Comparison of a traditional disease-specific and a quality-of-life measurement of outcome. J Arthroplasty 1997;12:639–45.

[16] Heck DA, Robinson RL, Partridge CM, Lubitz RM, Freund DA. Patient outcomes after knee replacement. Clin Orthop 1998;93–110.

[17] Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989;27:S217–32.

[18] Pouchot J, Guillemin F, Coste J, Bregeon C, Sany J. Validity, reliability, and sensitivity to change of a French version of the arthritis impact measurement scales 2 (AIMS2) in patients with rheumatoid arthritis treated with methotrexate. J Rheumatol 1996;23:52–60.

[19] Guillemin F, Coste J, Pouchot J, Ghezail M, Bregeon C, Sany J. The AIMS2-SF: a short form of the Arthritis Impact Measurement Scales 2. French Quality of Life in Rheumatology Group. Arthritis Rheum 1997;40:1267–74.

[20] Ren XS, Kazis L, Meenan RF. Short-form Arthritis Impact Measurement Scales 2: tests of reliability and validity among patients with osteoarthritis. Arthritis Care Res 1999;12:163–71.

[21] Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. Arthritis Rheum 1992;35:1–10.

[22] Fielden JM, Gander PH, Horne JG, Lewer BM, Green RM, Devane PA. An assessment of sleep disturbance in patients before and after total hip arthroplasty. J Arthroplasty 2003;18:371–6.

[23] Fitzgerald JD, Orav EJ, Lee TH, Marcantonio ER, Poss R, Goldman L, Mangione CM. Patient quality of life during the 12 months following joint replacement surgery. Arthritis Rheum 2004;51:100–9.

[24] The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995;41:1403–9.

[25] Stucki G, Sigl T. Assessment of the impact of disease on the individual. Best Pract Res Clin Rheumatol 2003;17:451–73.

[26] Stucki G, Ewert T, Cieza A. Value and application of the ICF in rehabilitation medicine. Disabil Rehabil 2002;24:932–8.

[27] Guillemin F, Coste J, Retel-Rude N, Spitz E, Baumann M, Tarquino C, Luttenbacher C, Pouchot J, and The French Group Quality of Life in Rheumatology. Strategy and methodology for choice of items in psychometric measurement: designing a quality of life instrument for of hip and knee osteoarthritis. In: Mesbah M, Cole BF, Ting Lee ML, editors. Statistical methods for quality of life studies: design, measurements and analysis. Dordrecht: Kluwer Academic Publishers; 2002. p. 51–61.

[28] Fisher RP, Falkner KL, Trevisan M, McCauley MR. Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. J Appl Psychol 2000;85:180–9.

[29] Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, Brown C, Cooke TD, Daniel W, Feldman D, Greenwald R, Hochberg M, Howell D, Ike R, Kapila P, Kaplan D, Koopman W, Marino C, McDonald E, McShane DJ, Medsger T, Michel B, Murphy WA, Osial T, Ramsey-Goldman R, Rothschild B, Wolfe E. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34:505–14.

[30] Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M, Howell D, Kaplan D, Koopman W, Longlay S III, Mankin H, McShane DJ, Medsger T, Meenan R, Mikkelsen W, Moskowitz R, Murphy W, Rothschild B, Segal M. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29:1039–49.

[31] Ware JE, Kosinski M, Keller SK. SF-36 physical and mental health summary scale: a user's manual. Boston, MA: The Health Institute; 1994.

[32] Streiner D, Norman G. Health measurement scales: a practical guide to their development and use. Ed 2. New York: Oxford University Press; 1995.

[33] Coste J, Fermanian J, Venot A. Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. Stat Med 1995;14:331–45.

[34] Shrout PE. Measurement reliability and agreement in psychiatry. Stat Methods Med Res 1998;7:301–17.

[35] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

[36] Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol 1998;51:945–52.

[37] Hays RD, Wang EYI. Multitrait scaling program: MULTI [RAND, University of Southern California] 2003. Available at: http://www.rand.org/publications/MR/MR162/.appc.pdf. Accessed July 2, 2003.

[38] McKenna SP, Hunt SM. A new measure of quality of life in depression: testing the reliability and construct validity of the QLDS. Health Policy 1992;22:321–30.

[39] Whalley D, McKenna SP, de Jong Z, van der Heijde D. Quality of life in rheumatoid arthritis. Br J Rheumatol 1997;36:884–8.

[40] Doward LC, Spoorenberg A, Cook SA, Whalley D, Helliwell PS, Kay LJ, McKenna SP, Tennant A, van der Heijde D, Chamberlain MA. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. Ann Rheum Dis 2003;62:20–6.

[41] Kaplan RM, Anderson JP, Wu AW, Mathews WC, Kozin F, Orenstein D. The Quality of Well-being Scale. Applications in AIDS, cystic fibrosis, and arthritis. Med Care 1989;27:S27–43.

[42] Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, Hiller TK. Evaluation of impairment of health-related quality of life in asthma: development of a questionnaire for use in clinical trials. Thorax 1992;47:76–83.

[43] Greene J, McClintock C. Triangulation in evaluation, design, and analysis issues. Evaluation Review 1985;9:523–45.

[44] Maggs-Rapport F. Combining methodological approaches in research: ethnography and interpretive phenomenology. J Adv Nurs 2000;31:219–25.

[45] Coste J, Guillemin F, Pouchot J, Fermanian J. Methodological approaches to shortening composite measurement scales. J Clin Epidemiol 1997;50:247–52.

[46] Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998;51:903–12.

[47] Leplege A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation, and preliminary psychometric evaluation. J Clin Epidemiol 1998;51:1013–23.

[48] Poiraudeau S, Chevalier X, Conrozier T, Flippo RM, Liote F, Noel E, Lefevre-Colau MM, Fermanian J, Revel M, Rhumato R. Reliability, validity, and sensitivity to change of the Cochin hand functional disability scale in hand osteoarthritis. Osteoarthritis Cartilage 2001;9:570–7.

[49] Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, William JI. The Quebec Back Pain Disability Scale. Measurement properties. Spine 1995;20:341–52.

[50] Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.

[51] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186–91.