A Composite Index for Total Hip Arthroplasty in Patients with Hip Osteoarthritis

JEAN FRANCIS MAILLEFERT, ALICE GUEGUEN, MINH NGUYEN, LAURENT BERDAH, MICHEL LEQUESNE, BERNARD MAZIERES, ERIC VIGNON, and MAXIME DOUGADOS

ABSTRACT. ***Objective.*** We propose a composite index for considering total hip arthroplasty (THA) in hip osteoarthritis (OA).

Methods. We carried out a 3 year longitudinal study of patients with painful hip OA from 137 centers. Clinical data were collected at baseline and every 3 months; radiographs were taken at entry and each year. The decision to have surgery was made by the patient, the rheumatologist, and the surgeon, with no reference to outcome measures. Statistical analysis included discrete Cox analysis with time dependent covariates, on 3 month interval grouped data. The dependent variable was THA during the 3 months following the evaluated visit. Time dependent covariates collated at each evaluated visit included radiological joint space width (JSW), percentage decrease in JSW during the year preceding, patient's global assessment, Lequesne index, pain, and nonsteroidal antiinflammatory drug and analgesic intake. We compared 2 analyses differing in symptomatic variables entered: values obtained at a single time point vs mean values between 2 visits at a 3 month interval. Selection of the index was based on the best combination of variables to predict occurrence of THA.

Results. Of the 508 patients recruited for study, 42 were excluded. During the 3 year followup, 75 patients underwent THA. Symptomatic, therapeutic, and radiological variables were included in the index resulting from the 2 analyses. Based on the selected cutoff, the positive and negative predictive values for occurrence of THA in the 2 years following were 54.3 and 90.6%, respectively (single point model), and 52.9 and 86.7%, respectively (model using mean values of symptomatic variables between 2 visits).

Conclusion. The poor positive predictive value of the composite indices obtained in this study suggests that there are other unmeasured factors determining access to surgery. On the other hand, the high negative predictive values suggest that these composite measures should be used by clinicians to determine which patients should not be referred to THA. (J Rheumatol 2002;29:347–52)

Key Indexing Terms:

TOTAL HIP ARTHROPLASTY	HIP OSTEOARTHRITIS	OUTCOME MEASURE
COMPOSITE INDEX		INDICATION

A composite index for considering total hip arthroplasty (THA) in hip osteoarthritis (OA) has several objectives: (1) to homogenize indications, which vary among centers and among surgeons[1-5]; (2) to better evaluate the need, and to prioritize the indications for THA, particularly in countries with long waiting periods for surgery; and (3) to help physicians to objectively determine the appropriate time for THA. For example, it has been recently suggested that there is underuse of THA for severe arthritis in Canada, especially in women[2]. Those investigators found the potential need for hip and knee arthroplasty was more than 3 times as great among women as among men[2]. Finally, it could be considered as a hard outcome measure in therapeutic trials. Besides symptomatic (e.g., pain, functional disability) and structural (e.g., degree of radiological joint space narrowing) variables[6-8], the occurrence of THA has been proposed as a clinically relevant outcome measure permitting evaluation of potential disease modifying treatments[9]. However, the wide international variability of the decision for surgery strongly suggests that the time to fulfil the criteria for considering THA instead of the time to surgery might be more appropriate as an endpoint. Such an index should have several components. The decision for THA in hip OA usually takes into account the following variables: (1) the symptomatic severity of the disease and its effect on lifestyle; (2) the structural severity of the disease; (3) other clinical elements, such as age and comorbidity; (4) the patient's willing-

ness to undergo THA; (5) the local health care system and the patient's status in it.

To our knowledge, 4 indices for THA have been proposed[2,10-12]: (1) the US National Institutes of Health suggested that THA should be considered in patients with radiographic evidence of joint damage, persistent pain, and disability interfering with daily activities, and that THA may not be recommended in patients who are at high risk for infections or in poor health[10]; (2) Hawker, *et al* defined potential candidates for THA as patients with a Western Ontario McMaster University Osteoarthritis Index (WOMAC) summary score ≥ 39, clinical and radiographic evidence of arthritis, and no absolute contraindication to THA; the estimate of potential need was then adjusted for willingness[2]; (3) Lequesne's index[11], currently used as an outcome measure in clinical trials, was designed as an index for considering THA; (4) the New Zealand score includes pain, functional impairment, range of motion, deformity, and other factors, such as effect on lifestyle[12]. However, these indices take into account neither the degree of severity of structural variables nor the therapeutic variables. Finally, to our knowledge, they have not been validated for sensitivity, specificity, and predictive values in large populations.

Our objective was to design and propose a data driven composite index for considering THA in hip OA, taking into account symptomatic, structural and therapeutic variables, and impact on lifestyle.

MATERIALS AND METHODS

Study design. A multicenter, prospective, longitudinal, 3 year followup study was designed. The study protocol was approved by the local ethics committee (Hôpital Cochin, Paris, France).

Inclusion criteria. These were as described[9]. Briefly, outpatients visiting a rheumatologist and fulfilling the American College of Rheumatology criteria for the diagnosis of hip OA[13] were enrolled after written informed consent was obtained. Other inclusion criteria were age between 50 and 75 years and hip pain on a daily basis for at least one month during the last 3 months. The exclusion criteria were radiological joint space width < 1 mm at the narrowest point, radiographic medial or axial femoral head migration, and secondary hip OA.

Outcome variables. Each patient was evaluated by a single rheumatologist at baseline, and every 3 months over 3 years.

Symptomatic outcome variables. At entry and at each evaluation, the following data were collected: pain occurring during physical activities during the previous 2 days [100 mm visual analog scale (VAS)], functional disability (Lequesne index, which is a 0–24 grade scale related to the effect of hip OA on daily activities), patient's overall assessment of disease activity during the previous 2 days (0–4 scale: none/mild/moderate/severe/very severe), nonsteroidal antiinflammatory (NSAID) and analgesic intake (number of days requiring such therapy).

Structural outcome variables. At entry and once a year, joint space width was evaluated in mm (using a 0.1 mm graduated magnifying glass) at the narrowest point on an anteroposterior weight bearing pelvic radiograph.

Decision for surgery. This was made by the patient, the rheumatologist, and the surgeon, with no reference to radiographic change and/or to analysis of objective symptomatic variables.

Statistical analysis. A discrete Cox model with time dependent covariates was performed on 3 month interval grouped data. The discrete Cox model is equivalent to a logistic regression model with complementary log-log transformation as link function[14]. Each discrete time unit for each subject is treated as a separate observation. For each observation, the dependent variable was occurrence of THA (yes or no) during the observed 3 months. To obtain an easy-to-use instrument, the time dependent categorical covariates were switched into a 3 grade scale based on distribution in the studied population (Table 1). Two analyses were performed. Structural variables were entered in both analyses (since 2 radiographs are required to calculate the decrease in joint space width, patients who underwent THA during the first year of followup were excluded), symptomatic variables were also entered in both analyses; however, in the first one the values were those obtained at the evaluated visit, i.e., at a single point of time, while in the other one values were those obtained at the last 2 visits, i.e., the mean value between the current visit and the one performed 3 months earlier.

The multivariate Cox models were performed with backward elimination of variables. The estimated parameter (EP) associated with each covariate was obtained (e^{EP} = relative risk for occurrence of THA during the 3 months following the evaluated visit). From each EP, a coefficient (EP × 100/maximal potential value of the sum of EP) was calculated and rounded off. A score was defined as the sum of the rounded off coefficients. The greater the score values, the higher the risk of occurrence of THA during the 3 months following the evaluated visit. For each cutoff of the score, the correct classification probabilities in observations with occurrence of THA or not were obtained. The best cutoff value was then determined from the graphic representation of correct classification probabilities.

Since numerous subjects were evaluated several times, the number of observations with no occurrence of THA during the 3 months following the evaluated visit was expected to exceed by far the number of observations with occurrence of THA. Consequently, the chosen cutoff scores were expected to have a high negative but a poor positive predictive value for occurrence of THA or not during the 3 months following all the evaluated visits. Thus, to evaluate the clinical usefulness of the proposed index, the positive and negative predictive values of the chosen cutoff scores were evaluated at a single visit, i.e., Month 12. For this purpose, the clinical, structural, and therapeutic variables obtained at this visit were collated, permitting calculation of the scores of the 2 proposed composite indices for each individual patient. Thereafter, the risk of occurrence of THA was evaluated first between Months 12 and 24, and second between Months 12 and 36.

RESULTS

Of the 508 patients recruited (205 men, 303 women; mean age 63 ± 7 yrs), 42 were excluded because THA was performed during the first year (39 patients) or because of missing data (3 patients). The main characteristics of the 466 patients are summarized in Table 2. During Years 2 and 3, THA was performed in 75 patients. Consequently, 3425 observations were analyzed.

Multivariate logistic regression analysis suggested that 5 variables (not identical in the 2 analyses) were determinants of occurrence of THA during the 3 months following the evaluated visit (Tables 3 and 4). These analyses permit easy calculation of a score for a single patient. For example, based on the first proposition (Table 3), a patient with a joint space width of 1.5 mm on the last available radiograph, who considers his/her disease activity as moderate, with a Lequesne index of 8, and without treatment, has a total score of 20 + 0 + 0 + 0 + 0 = 20. From the graphic representations (Figure 1), the best cutoff value of the score was 40 (< 40 and ≥ 40) in the 2 analyses. In the analysis using symptomatic and therapeutic variables obtained at the evaluated visit the sensitivity and specificity were 82.7 and 84.1%, respectively. In the second analy-

Table 1. Symptomatic radiological and therapeutic variables entered in the analyses. The cutoff points were chosen at the approximate 33rd and 66th percentiles of the population at inclusion.

Time Dependent Covariates	3 Grade Scale		
Joint space width (last radiograph available), mm	> 2	≤ 2 and > 1	≤ 1
Percentage decrease in joint space width between the 2 last annual radiographs	< 10	> 10 and ≤ 50	> 50
Pain occurring after physical activities during the 2 previous days (100 mm VAS) of the evaluated visit	≤ 40	> 40 and ≤ 60	> 60
Patient's overall assessment of disease activity during the previous 2 days (0–4 scale; none/mild/moderate/severe/very severe) of the evaluated visit	≤ 2	> 2 and ≤ 3	> 3
Lequesne index considering the previous 2 days of the evaluated visit	≤ 10	> 10 and ≤ 12	> 12
NSAID intake during the 3 mo prior to the evaluated visit	≤ 1 day/week	> 1 day/week and ≤ 1 day/2	> 1 day/2
Analgesic intake during the 3 mo prior to the evaluated visit	≤ 1 day/week	> 1 day/week and ≤1 day/2	> 1 day/2

Table 2. Baseline characteristics of the 466 patients.

Variables	Results
Demographic data	
Age, yrs (mean ± SD)	63 ± 6.8
M/F, no.	194/272
Body mass index, kg/m^2 (mean ± SD)	25.8 ± 3.5
Symptomatic data,	
Pain, mm (mean ± SD)	43.6 ± 19.9
Lequesne index (mean ± SD)	7.6 ± 2.5
Patient overall assessment, no. of patients	
0 (none)	9
1 (mild)	105
2 (moderate)	248
3 (severe)	99
4 (very severe)	5
Radiographic data, no. of patients	
Femoral head migration	
Superolateral	268
Superomedial	154
Concentric	44
Joint space width, mm (mean ± SD)	2.31 ± 0.83

sis, the sensitivity and specificity were similar, 85.3 and 80.1%, respectively.

Positive and negative predictive values of the chosen cutoff scores for occurrence of THA or not during the one and 2 years following the visit performed at Year 1 are shown in Table 5.

Finally, to evaluate whether the 39 patients with THA during the first year differed or not from other patients, their score obtained at the last visit before THA was calculated. The index using the symptomatic and therapeutic variables collated at a single visit, together with the radiological variables, was used at this step (the other index needs 2 annual radiographs, thus could not be obtained). In this population, the mean score obtained at the last visit before THA was 50.4 ± 24.5. A value > 40 was observed in 27 patients.

DISCUSSION

Two composite indices for considering THA in hip OA are proposed. Several factors suggest that they might be relevant in both daily practice and clinical research. The variables

Table 3. Results of multivariate logistic regression in which the occurrence of THA was the dependent variable and the independent variables were symptomatic and therapeutic variables collected at a single visit together with the radiological variables.

Time Dependent Covariates	Grades of the Covariates	Relative Risk*	95% CI	Coefficient**	Rounded off Coefficient***
Joint space width on the last	> 2	1		0	0
available radiograph, mm	≤ 2 and > 1	3.9	1.8–9.16	20.16	20
	≤ 1	11.68	5.78–26.28	36.4	35
Patient's overall assessment					
(none/mild/moderate/severe/very severe)	≤ 2	1		0	0
(value collated at the evaluated visit)	> 2	2.85	1.53–5.32	15.52	15
Lesquesne index (value collated at	≤ 12	1		0	0
the evaluated visit)	> 12	5.62	3.13–10.15	25.57	25
NSAID intake during the 3 mo					
preceding the evaluated visit	≤ 1 day/2	1		0	0
	> 1 day/2	2.31	1.34– 3.94	12.4	15
Analgesic intake during the 3 mo	≤ 1 day/week	1		0	0
preceding the evaluated visit	> 1 day/week	1.98	1.16–3.4	10.1	10

* Relative risk for occurrence of THA during the 3 months following the evaluated visit. ** Coefficient = EP × 100/maximal potential value of the sum of EP, in which EP = Ln of relative risk for occurrence of THA during the 3 months following the evaluated visit. EP: estimated parameter. *** Coefficients were rounded off in order to obtain an easy-to-use instrument.

Table 4. Results of multivariate logistic regression in which the occurrence ot THA was the dependent variable, and the symptomatic and therapeutic variables collected at the 2 last visits, i.e., the mean value between the evaluated visit and the one performed 3 months ago, together with the radiological variables, were the independent variables.

Time Dependent Covariates		Relative Risk*	95% CI	Coefficient**	Rounded off Coefficient***
Joint space width on the last	> 2	1		0	0
available radiograph, mm	≤ 2 and > 1	3.37	1.55–7.9	18.5	20
	≤ 1	6.73	2.99–16.29	29.1	30
Percentage decrease in joint space width	< 50	1		0	0
between the 2 last annual radiographs	≥ 50	1.87	1.02–3.46	9.56	10
Mean value of patient's overall	≤ 2	1		0	0
assessment (none/mild/moderate/severe/	3 ≥ 2	3.42	1.8–6.58	18.75	20
very severe) (values collated at the 2 last	> 3	8.17	3.7–17.96	32.02	30
evaluated visits)					
Mean value of Lequesne index (values	≤ 10	1		0	0
collated at the 2 last evaluated visits)	> 10	3.1	1.77–5.47	17.24	20
Mean NSAID intake during the					
3 mo preceding the 2 last evaluated	≤ 1 day/2	1		0	0
visits (i.e., the 6 mo preceding the	> 1 day/2	2.21	1.28–3.78	12.12	10
last evaluated visit)					

* Relative risk for occurrence of THA during the 3 months following the last evaluated visit. ** Coefficient = EP × 100/maximal potential value of the sum of EP, in which EP = Ln of relative risk for occurrence of THA during the 3 months following the evaluated visit. EP: estimated parameter. *** Coefficients were rounded off in order to obtain an easy-to-use instrument.

included in the proposed design are valid and reliable[4,6-8,15]. The indices take into account the symptoms, effect on lifestyle (through the patient's assessment), structure, and concomitant required treatment. The variables are easy to collect (limited to 5, switched to a 2 or 3 grade scale, coefficients rounded off, easy to total). In this study, the symptomatic severity of hip OA has been assessed by evaluating several domains (pain using a 100 mm VAS, patient's overall assessment of disease activity using a 0–4 grade scale, functional impairment using the Lequesne index). All these domains and the tools used for evaluation have been recommended by the OMERACT[16] and the OARSI[17] societies. Both the patient's overall assessment and the Lequesne index have been selected in the proposed

indices. Such tools are widely used in the medical rheumatologic literature, but are not so used in the orthopedic literature. It might be interesting to conduct other studies using other tools such as the WOMAC[18], the Oxford hip score[19], and the Patient-Specific Index[20].

Two indices are proposed, differing in how the symptomatic and therapeutic variables are collected. The analysis using collection of variables at a single point in time is easier to use than the second one, which requires 2 evaluations at a 3 month interval. However, the advantage of the second set is that, theoretically, it is insensitive to fluctuations in disease activity over time. The similarity in performance between the 2 methods may be due to the fact that both take into account

Table 5. Evaluation, for each proposition, of the positive and negative predictive values of the chosen cutoff scores, at a single visit, i.e., Month 12. The clinical, structural, and therapeutic variables obtained at this visit were collected, permitting calculation of each patient's score, and thereafter, the risk of occurrence of THA was evaluated first between Months 12 and 24, and second between Months 12 and 36.

Composite Index	Patients with a Score ≥ 40 (N)	Patients with a Score ≥ 40 with THA During the Year Following the Evaluated Visit, %	Patients with a Score ≥ 40 with THA During the 2 Years Following the Evaluated Visit, %	Patients with a Score < 40 (N)	Patients with a Score < 40 with no THA During the Year Following the Evaluated Visit, %	Patients with a Score < 40 with no THA During the 2 Years Following the Evaluated Visit, %
First proposition*	70	38.6	54.3	396	96.2	90.6
Second proposition **	68	41.2	52.9	398	96.5	86.7

* Index using the symptomatic and therapeutic variables collated at a single visit, together with the radiological variables (Table 3). ** Index using the symptomatic and therapeutic variables collated at the 2 last visits, i.e., the mean value between the evaluated visit and the one performed 3 monthes earlier, together with the radiological variables (Table 4).
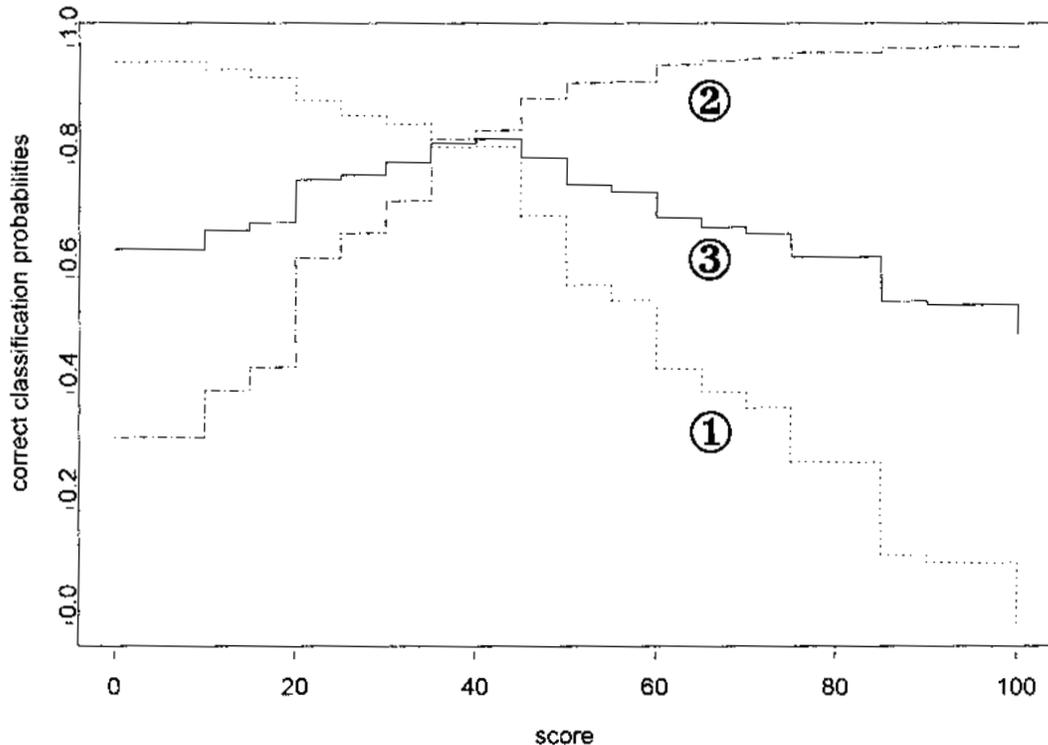
Figure 1. Analysis using symptomatic and therapeutic variables obtained at the last evaluated visit. Graphic representation of correct classification probabilities. The best cutoff value of the score was 40 (< 40 and ≥ 40), with sensitivity and specificity of 82.7 and 84.1%, respectively. In the analysis using symptomatic and therapeutic variables obtained at the 2 last evaluated visits, the graphic representation of correct classification probabilities was similar (data not shown). 1: percentage of correct classification probabilities for observations going to THA in the 3 months following; 2: percentage of correct classification probabilities for observations not going to THA in the 3 months following; 3: mean percentage of correct classification probabilities.

the persistence of symptomatic severity, since the therapeutic variable (daily intake of NSAID or analgesic) reflects disease severity during the 3 months preceding the visit.

Several criticisms of this proposal could be made. First, although the number of patients included was high, the population might not be representative of the general population of patients with hip OA. In particular, patients with painless hip OA and those at an advanced radiological stage were excluded. However, THA is rarely considered in patients without pain, and is readily considered in patients at an advanced stage of the disease.

Second, one could argue that the proposed indices do not indicate which patients require THA, but in which patients THA was performed. Patients were followed by rheumatologists. The decision for THA was made by the patient and by specialized practitioners (the rheumatologist and the surgeon). Moreover, since the decision took into account the opinion of the patient and those of at least 2 physicians, and since the study was a multicenter design, some disparities in surgical decisions should have been taken into account. Moreover, we reported previously that the percentage of patients requiring hip surgery was not statistically different with regard to residence[9]. In addition, economic considerations should not have influenced the results, since the great majority of the French population is insured by the National Health Service. In particular, in France, there is no delay due to the health care system between the time of the visit during which the decision to undergo surgery is made and the date of the surgical intervention.

Finally, the positive predictive values of the cutoff scores for occurrence of THA or not in the subsequent one and 2 years could be considered unsatisfactory. Nearly one-half of the patients with a score ≥ 40 did not undergo THA within the 2 following years. These results could reflect the potential underuse of surgery in hip OA. They also could be explained by the fact that the other components included in the decision for surgery were not analyzed in this study, i.e., age, comorbidity, and patient's willingness. In spite of this, these indices could represent the first step in decision making for surgery. In particular, the discordance between poor positive predictive values and high negative predictive values suggests that the proposed composite measures should be used by clinicians to determine which patient should not be referred to THA. In clinical practice, it might provide information regarding the severity of OA, and makes it possible to determine the time to surgery on the basis of an objective instrument. Finally, it must be pointed out that the choice of the cutoff score depends on the goals of the indices. If the cutoff is decreased, then the

positive predictive value decreases, and the negative predictive value increases. For example, if the goal is to increase referral consideration of patients for THA, one might prefer an index with a lower positive predictive value (some individuals who do not need THA will be referred, but few who need it will be missed).

These composite indices should be regarded as preliminary and should be validated in further studies conducted with different aims (clinical trial vs daily practice vs prioritization of waiting list) with different characteristics of patients and in different countries with different health care systems.

REFERENCES

1. Charlson ME, Allegrante JP. Disparities in the use of total joint arthroplasty. N Engl J Med 2000;342:1044-5.
2. Hawker GA, Wright JG, Coyte PC, et al. Differences between men and women in the rate of use of hip and knee arthroplasty. N Engl J Med 2000;342:1016-22.
3. Mancuso CA, Ranawat CS, Esdaile JM, Johanson NA, Charlson ME. Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. J Arthroplasty 1996;11:34-46.
4. Dieppe P, Brandt KD, Lohmander S, Felson DT. Detecting and measuring disease modification in osteoarthritis. The need for standardized methodology. J Rheumatol 1995;22:201-3.
5. Fortin PR, Clarke AE, Joseph L, et al. Outcomes of total hip and knee replacement. Preoperative functional status predicts outcomes at 6 months after surgery. Arthritis Rheum 1999;42:1722-8.
6. Lequesne M, Brandt K, Bellamy N, Moskowitz R, Menkes CJ, Pelletier JP. Guidelines for testing slow acting drugs in osteoarthritis. J Rheumatol 1994;21:65-73.
7. Group for the Respect of Ethics and Excellence in Science (GREES). Recommendations for the registration of drugs used in the treatment of osteoarthritis. Ann Rheum Dis 1996;55:552-7.
8. Altman RD, Brandt KD, Hochberg MC, Moskowitz RM, for the Task force of the Osteoarthritis Research Society. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a Task Force of the Osteoarthritis Research Society. Osteoarthritis Cartilage 1996;4:217-43.
9. Dougados M, Gueguen A, Nguyen M, et al. Requirement for total hip arthroplasty: an outcome measure of hip osteoarthritis. J Rheumatol 1999;26:855-61.
10. NIH consensus development panel on total hip replacement. JAMA 1995;273:1950-6.
11. Lequesne M. The algofunctional indices for hip and knee osteoarthritis. J Rheumatol 1997;24:779-81.
12. Hadorn HC, Holmes AC. Education and debate: the New Zealand priority criteria project, part I: overview. BMJ 1997;314:131-4.
13. Altman R, Alarcon G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34:505-14.
14. Hosner DW, Lemeshow S. Logistic regression analysis of survival data. In: Hosner DW, Lemeshow S, editors. Applied logistic regression. New York: John Wiley & Sons; 1998:238-45.
15. March LM, Brooks PM. Clinical trials in osteoarthritis. Ann Rheum Dis 1996;55:491-3.
16. Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24:799-802.
17. Hochberg MC, Altman RD, Brandt KD, Moscowitz RW. Design and conduct of clinical trials in osteoarthritis: preliminary recommendations from a task force of the Osteoarthritis Research Society. J Rheumatol 1997;24:792-4.
18. 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.
19. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg 1996;78B:185-90.
20. Wright JG, Young NL. The patient-specific index: asking patients what they want. J Bone Joint Surg 1997;79A:974-83.