

# The Oxford hip and knee outcome questionnaires for arthroplasty

OUTCOMES AND STANDARDS FOR SURGICAL AUDIT

# P. B. Pynsent, D. J. Adams, S. P. Disney

From the Royal Orthopaedic Hospital, Birmingham, England The Oxford hip and knee scores are used to measure the outcome after primary total hip and knee replacement. We propose a new layout for the instrument in which patients are always asked about both limbs. In addition, we have defined an alternative scoring method which accounts for missing data. Over a period of 4.5 years, 4086 (1423 patients) and 5708 (1458 patients) questionnaires were completed for hips and knees, respectively. The hip score had a pre-operative median of 70.8 (interquartile range (IQR) 58.3 to 81.2) decreasing to 20.8 (IQR 10.4 to 35.4) after one year. The knee score had a pre-operative median of 68.8 (IQR 56.2 to 79.2) decreasing to 29.2 (IQR 14.6 to 45.8). There was no further significant change in either score after one year. As a result of the data analysis, we suggest that the score percentiles can be used as a standard for auditing patients before and after operation.

The need to measure the outcome of surgery has been recognised for many years. However, it is only in the last decade that the routine use of outcome instruments has been introduced into orthopaedic practice. There are several measures available for total hip (THR) and total knee (TKR) replacements but the Harris<sup>1</sup> and Charnley<sup>2</sup> hip scores, and the Hospital for Special Surgery knee score, have dominated the literature.<sup>3,4</sup> The traditional approach has been to measure both signs and symptoms, a system which imposes a heavy burden on clinical resources. As a consequence, the emphasis has moved to the design of patient self-reporting questionnaires which are simpler to implement and have improved because of the increased psychometric knowledge on validation and reliability which is now available.<sup>5</sup>

Despite the availability of outcome instruments, they are infrequently used for routine surgical audit. In this prospective study we present the results of applying Oxford hip and knee questionnaires<sup>6,7</sup> to patients undergoing THR and TKR, with their four-year follow-up results. Some standards for the use of these scores in auditing the progress of patients are given. Changes made to the layout of the hardcopy version of these instruments, and to the scoring method, are also presented.

## **Patients and Methods**

Questionnaires were given to patients over the age of 16 years who underwent primary THR or TKR at the Royal Orthopaedic Hospital,

Birmingham. The knee measure was introduced in 1997, initially for a single surgical firm, so that any logistical problems could be resolved. Gradually, over a year, its use spread to the entire arthroplasty service of the hospital. At the end of the first year, the hip questionnaire was introduced. Patients who underwent a joint replacement for tumours were excluded from the analysis. Our results include data to July 31, 2003. For pre-operative data, only measurements made at six months or less before the operation were included. The hospital undertook almost 2000 primary joint replacements each year over the period of study. Given this large number and limited resources, it was impossible to recruit every eligible patient. Selection of patients' was not systematic but consisted of recruiting and gaining consent from as many patients as possible each day from the outpatient clinic. It is possible that this method may have introduced bias into our results.

**Instruments.** The Oxford hip and knee scores are higher when the disability is greater. In the original publications the layout and scoring of the two questionnaires were very similar.<sup>6,7</sup> They were designed for a single joint and comprise 12 questions. Each has a stem and five graded alternatives which are scored from one to five (Table I).

The layout of the questionnaires has been changed (Fig. 1). We found that patients commonly had problems with both limbs and found it difficult to answer the questions on a

 P. B. Pynsent, PhD, Director
D. J. Adams, MA, BSc, RGN, Research Nurse
S. P. Disney, BSc, Clinical Scientist
Research and Teaching
Centre, Royal Orthopaedic
Hospital, Birmingham B31
2AP, UK.

Correspondence should be sent to Dr P. B. Pynsent; e-mail:

p.b.pynsent@bham.ac.uk

©2005 British Editorial Society of Bone and Joint Surgery doi:10.1302/0301-620X.87B2. 15095 \$2.00

J Bone Joint Surg [Br] 2005;87-B:241-8. Received 5 November 2003; Accepted after revision 9 June 2004

#### Table I. The stem portion of the Oxford hip and knee questionnaires<sup>6,7</sup>

THR	
1	How would you describe the pain you <u>usually</u> had from your hip?
2	Have you had any trouble with washing and drying yourself (all over) because of your hip?
3	Have you had any trouble getting in and out of a car or using public transport because of your hip? (whichever you tend to use
4	Have you been able to put on a pair of socks or tights?
5	Could you do the household shopping on your own?
6	For how long have you been able to walk before <u>pain from your hip</u> becomes severe?
7	Have you been able to climb a flight of stairs?
8	After sitting at a table (i.e. for a meal), how painful has it been for you to stand up from a chair <u>because of your hip</u> ?
9	Have you been limping when walking <u>because of your hip</u> ?
10	Have you had any sudden, severe pain - 'shooting', 'stabbing' or 'spasms' - <u>from the affected hip</u> ?
11	How much has <u>pain from your hip</u> interfered with your usual work (including housework)?
12	Have you been troubled by <u>pain from your hip</u> in bed at night?
TKR	
1	How would you describe the pain you <u>usually</u> had from your knee?
2	Have you had any trouble with washing and drying yourself (all over) <u>because of your knee?</u>
3	Have you had any trouble getting in and out of a car or using public transport because of your knee? (whichever you tend to use
4	For how long have you been able to walk before <u>pain from your knee</u> becomes severe?
5	After sitting at a table (i.e. for a meal), how painful has it been for you to stand up from a chair because of your knee?
6	Have you been limping when walking, because of your knee?
7	Could you kneel down and get up again afterwards?
8	Have you been troubled by <u>pain from your knee</u> in bed at night?
9	How much has <u>pain from your knee</u> interfered with your usual work (including housework)?
10	Have you felt that your knee might suddenly 'give way' or let you down?
11	Could you do the household shopping on your own?
12	Could you walk down one flight of stairs?

unilateral basis. We therefore adapted the questionnaires so that pain and function could be recorded bilaterally. This allowed the patients to apportion disability to a specific side and extended the scope of the instrument, not only for comparison between sides but also for the observation of changes in the contralateral joint.

The scoring of the questionnaires was also changed. The published method was cumbersome. A healthy joint would

1	During the past 4 weeks								
	How would you describe the pain you <u>usually</u> had from your hip?								
Left Right	None [] []	Very mild	Mild [] []	Moderate [] []	Severe [] []				
2	2 During the past 4 weeks Have you had any trouble with washing and drying yourself (all over) because of your hip?								
Left Right	No trouble at all [] []	Very little trouble	Moderate trouble [] []	Extreme difficulty [] []	Impossible to do [] []				
3	<i>During the past 4 weeks</i> Have you had any trouble getting in and out of a car or using public transport <u>because of your hip?</u> (whichever you tend to use)								
Left Right	No trouble at Very little all trouble		Moderate trouble ] ]	Extreme difficulty [] []	Impossible to do [] []				
4	4 During the past 4 weeks Have you been able to put on a pair of socks, stockings or tights? Yes, easily With little With moderate With No, difficulty difficulty extreme impossible								
Left Right									

#### Fig. 1

Part of the modified layout of the hip questionnaire showing how each question is asked separately for both left and right sides.



The box-and-whisker plot is constructed with a filled bar (i.e. box), the ends of which represent the 25% and 75% percentiles. The horizontal line within the bar is the median. The difference between the 25th and 75th percentiles is the interquartile range (IQR). The indents on the bar sides represent the approximate 95% Cls for the median. If, in any two boxes, these notches do not overlap, then the medians are significantly different at the 5% level. The upper whisker extends to the most extreme data point which is no more than 1.5 times the IQR from the 75th percentile. Similarly, the lower whisker extends to the most extreme data point which is not less than 1.5 times the IQR from the 25th percentile. The upper whisker extends to the most extreme data point which is not less than 1.5 times the IQR from the 25th percentile. The mean (diamond symbol) and its 95% Cls (arrows) are superimposed in Figures 3 and 4.

gain a final score of 12 with 12 questions scoring one point each and the worst possible joint would gain a score of 60  $(12 \times 5)$ . We adopted an alternative scoring system in which each question was scored between zero and four and the final index was expressed as a percentage. A healthy joint thus scored 0% and the worst possible joint 100%. In the original publications no indications were given as to how to record a score if the questionnaire was incomplete or if multiple boxes had been marked for one question. This could lead to ambiguity. For example, a totally healthy patient who omitted one question would score 11 points. We therefore used the system defined for the Oswestry Disability Questionnaire.<sup>8</sup> If multiple boxes were marked, the highest mark was used. For omitted questions, the score was expressed as a percentage of the total score possible, but only using those questions which were actually answered. Any questionnaire with more than two omitted questions was excluded.

**Storage and analysis of data**. Data were stored on the hospital's audit database. This is an SQL database (Oracle Corporation, Redwood Shores, California) which is connected to the hospital's intranet. In addition to outcomes, the data-

VOL. 87-B, No. 2, FEBRUARY 2005

base stores other information such as details of the patient, waiting times, operations, health status and complications. Our results have been taken from their database and were analysed using the R statistical package.<sup>9</sup> Many of the results are presented as box-and-whisker plots. An explanation of how to interpret these is given in Figure 2. Interquartile ranges (IQR) have been expressed as ranges between limits rather than the absolute differences in order to emphasise asymmetry around the median. We have also quoted the 95% confidence interval (CI).

Acquisition of data. According to the hospital's protocol patients undergoing THR and TKR should be measured pre-operatively and at six weeks, three months (knees only) and one year post-operatively at follow-up clinics. Subsequently, they should be sent the questionnaire annually by post. In practice, the follow-up times varied substantially from these planned intervals. Our results are therefore grouped into six-month periods rather than fixed-time intervals. Some of the data were obtained directly from computerised records using in-house software which had been designed for patient interviews.<sup>10,11</sup> This facility was withdrawn by the hospital management at an early stage of

244

Table II. Details of hips and knees in the study, by number and *percent-age*. The diagnoses and hospital percentages have been obtained from the hospital's computerised records and not the study database. For 26% of THRs and 31% of TKRs the diagnoses were not coded

	THR	TKR
Number of joints (%)		
Male	534 <i>(34)</i>	636 <i>(37)</i>
Female	1020 <i>(66)</i>	1103 <i>(63)</i>
Total	1554	1739
Sample of total in hospital (%)	19	17
Mean age in years (range)		
Men	67 (19 to 92)	70 (18 to 92)
Women	70 (17 to 94)	70 (21 to 94)
Total	60 (SD 11.4)	70 (sd 10)
Diagnosis (%)		
Osteoarthritis	91	91
Rheumatoid arthritis	3	7
Other	6	2

our study so that most questionnaires (approximately 95%) were completed in their hard-copy format either in the clinic or by postal follow-up. The answers from this hard-copy version were transcribed to the database by clerical staff.

## Results

Over the period of study 4086 THR questionnaires (1423 patients, 131 bilateral) met the criteria for analysis. Ten questions were answered in 68 (2%) and 11 in 291 (7%) questionnaires. For TKR there were 5708 (1458 patients, 281 bilateral) responses. This sample represents 19% of the THRs undertaken and 17% of the TKRs undertaken in the hospital during the period of the study. The distribution of these patients in the two groups, and some of their details, are shown in Table II. The number of forms and the mean follow-up times are summarised in Table III. For the hips, of the 4086 forms, 1663 (41%) questionnaires were completed for the operated side only, 483 (12%) for both sides and 1940 (47%) discriminated between the two sides. For the knees, of the 5708 forms, 2038 (36%) questionnaires were completed for the operated side only while 690(12%)were completed identically for both knees. This left 2980 (52%) questionnaires which discriminated between the two knees.

The distribution of the pre- and post-operative scores for the hips and knees is shown in Figure 3. The number of forms completed and the mid-point for each six-month interval shown in Figure 3 are given in Table III. The hip plot showed a steady and significant decline in the median score, judged by the non-overlapping CIs for the first postoperative year, from a pre-operative value of 70.8 (IQR 58.3 to 81.2) to 20.8 (IQR 10.4 to 35.4). There was then no further significant change in the median for the remainder of the period of study. For the knees, again there was a significant decline in the median score for the first year, from 68.8 (IQR 56.2 to 79.2) to 29.2 (IQR 14.6 to 45.8) after operation and then there was no further significant change. The way in which the score distribution for individual answers had changed to account for these lower post-operative values is illustrated in Figure 4 for the 1.5- to two-year interval. For example, for the data on the knees it can be seen that the pre- and post-operative median for question 7 remains at a maximum score of four points. Thus the patients cannot kneel before the operation and still cannot kneel afterwards.

## Discussion

Since publication, the Oxford hip and knee questionnaires have been used by several authors for measuring self-reported disabilities. Most of these publications are from the UK<sup>12,13</sup> with the North American literature showing a preference for the WOMAC index.<sup>14-16</sup> Dunbar et al<sup>17</sup> found a better response rate for the Swedish version of the Oxford knee questionnaire compared with the WOMAC index. The Oxford instruments were chosen for our study because they are valid,<sup>18</sup> reliable,<sup>6,7</sup> free and the designs are based on the Oswestry disability index which has been successful in assessing spinal problems.<sup>8,19</sup>

The changing of any outcome instrument is fraught with problems.<sup>5</sup> Although the wording of the actual questions in the instruments has not been changed, there is no reason to suggest that the validity of the questions has altered. The application of the instrument to both sides simultaneously may produce different results. In a set of single-sided applications of the Oxford hip questionnaire, patients were also asked to comment freely on their operation.<sup>20</sup> Some with bilateral disease scored a nearly maximum disability but commented that their operation was "wonderful and they could not wait for the other side to be done". This clearly was not the required response. By contrast, a question such

Table III. The mean time-values and number of questionnaires for the hips and knees. This table relates to the box-and-whisker plots of Figure 3 (i.e. each box does not exactly fall on the six-month time interval shown on the x-axis). Negative times up to and including zero represent the pre-operative time

	Time interval centres (yrs)									
	Pre-operative	0.25	0.75	1.25	1.75	2.25	2.75	3.25	3.75	4.25
Knee forms										
Number	1438	1881	716	571	265	302	210	158	90	77
Mean time in years (SD)	-0.09 (0.10)	0.22 (0.13)	0.77 (0.15)	1.19 (0.14)	1.75 (0.15)	2.25 (0.14)	2.76 (0.14)	3.24 (0.14)	3.75 (0.15)	4.19 (0.13)
Hip forms										
Number	1341	1185	319	447	198	275	126	89	53	53
Mean time in years (SD)	-0.09 (0.09)	0.16 (0.09)	0.83 (0.14)	1.19 (0.14)	1.76 (0.15)	2.22 (0.14)	2.77 (0.14)	3.23 (0.14)	3.75 (0.15)	4.22 (0.14)



Box-and-whisker plots showing the distribution of the scores for a) hips and b) knees relative to the date of operation. The scores have been grouped into six-monthly intervals with the x-axis showing the mid-point of each interval. The left-hand y-axis shows our scoring system and the right-hand axis shows the approximate score of the original publications.

as "Have you been able to climb a flight of stairs?" is unambiguous since it either can or cannot be accomplished to some extent. Despite this, many patients answer the questions in such a way as to attribute all their limitations of activity and pain to one side of their body. McGrory and Harris<sup>21</sup> used the WOMAC index (disease-specific) and a self-reporting adaptation of the Harris hip score (joint-specific) to investigate if the WOMAC index could be used for bilateral disease. They used correlation to measure agreement and did not specify how the questionnaires were applied. They concluded that the WOMAC index could be used to measure disability in a bilateral situation. McMurray et al<sup>22</sup> commented on the problems of bilateral disease and other comorbidities in a review of the Oxford hip score. Harcourt, White and Jones<sup>23</sup> drew attention to the problem of specificity of the knee score when lumbar spine and hip morbidity co-existed. It may therefore be that scores from the new layout of the questionnaire (Fig. 1) cannot be compared directly with those obtained from the original layout.



Box-and-whisker plots showing the distribution of 12 questions scores pre-operatively (dark shade) and at the 1.5 to two-year time period (light shade) to show the change in individual questions for a) hip and b) knee questionnaires. When the Cl exceeds the IQR, the notch is extended beyond the box. See the legend for Figure 3 for an explanation of the two y-axes.

The justification for changing the scoring systems was based upon two factors. First, the original scoring system was not intuitive. We felt that it was unnatural to have an index for which a healthy joint scores 12 and for a completely disabled joint scores 60. We considered it to be more appropriate to have a score of 0% and 100%, respectively. Secondly, the method of scoring was not well defined since the original hip and knee questionnaires did not describe how to deal with missing data. A healthy joint should score 12 points in the original system but a missing answer would change the score to 11 points. Our method gave a score of 0% in both cases. The management of incomplete responses or of multiple responses to one question is most important. This problem always exists when using hard-copy questionnaires. For example, Robertsson and Dunbar<sup>24</sup> reported 10.6% of incomplete questionnaires in 1194 patients with the knee score. Fitzpatrick et al<sup>25</sup> reported 8.1%, 11.7% and 12.7% of incomplete questionnaires in the pre-operative, threemonth and 12-month applications of the hip score. The proposed change of the scale is linear if all questions are answered, that is:



Line plots showing the percentiles for a) hips and b) knees. See Figure 3 for an explanation of the two y-axes.

## y = (100 - 0)/(60 - 12) x - 25

where x is the original scoring system, y is the new percentage system and the slope is given by the ratio of the maximum range of score in each scale. This change in scale is analogous to changing from Fahrenheit to Celsius on a thermometer. Any inferences made do not change as long as the scale is linear. If only 11 or ten questions are answered, the slopes become 100/(55 - 11) and 100/(50 - 10), respectively, and the theoretical intercept stays constant at -25. The missing data in our study are small so that a best-fit straight line through the data has a minimal impact on the results (slope  $\neq 2.08$  and 2.08. intercept  $\neq -24.45$  and -24.33 for hips and knees, respectively). If data are reported without accounting for missing or mismarked questions, the results are impossible to interpret.

The results of Figure 3 confirm the success of these two operations in reducing pain and disability. In a large study (n = 7151), Fizpatrick et al<sup>25</sup> analysed the results for hip replacement over a 12-month period and reported a preoperative mean score of 67.7 (95% CI 67.3 to 68.1 when converted to new units) and a three-month follow-up mean score of 28.5 (95% CI 28.1 to 29.0). These means and intervals are within the 95% CI for the mean value which we obtained (Fig. 3b). However, because of the asymmetry of the data, the mean was not a good value to use in our study. Comparative results for the same time period as our study are not available, although Dawson et al,<sup>26</sup> in a study which compared two hip arthroplasties at seven years, reported median values of 30% and 35% ( $n_{total} = 331$ , both converted to new units) for Charnley (Johnson & Johnson Medical Ltd, Ascot, UK) and Hi-nek (Corin Medical, Cirencester, UK) prostheses, respectively.

The high post-operative values of question 7 (the ability to kneel) of the knee instrument (Fig. 4a) are important. Palmer et al<sup>27</sup> has noted that many patients only perceive that they cannot or should not kneel. This question partly contributes to the higher steady-state score for knees. The low pre-operative median score for question 2 (washing), in both the hip and knee questionnaires, suggests that this activity does not pose a particularly serious problem for patients, although the median score does significantly change after surgery to 1 (THR) and 0 (TKR). The sensitivity to change of the Oxford scores was established at the time of its first publication, and subsequently by the original authors.<sup>28</sup>

The data presented in Figure 3 are not useful as a reference standard for audit. The ideal would be to produce a standard chart, similar to a paediatric growth chart, by which patients can easily be monitored at outpatient clinics. Figure 5 shows such charts for the hip and knee produced by best-fit curves to the percentiles of Figure 3 using a simple exponential model. These data can be transformed to straight-line plots by using a log scale for the y-axis for those who prefer such a presentation. These plots have not been constructed from longitudinal data and represent values which could be expected with a single application of the instrument. They have not been constructed from information about patients' changes in score with time. From Figure 5, an individual surgeon can determine the degree of audit which should be introduced. Thus, it may be decided that a patient who was below the 25% percentile pre-operatively should be checked to see why, with such a low score, an operation was planned. Similarly, post-operatively it may be decided that any patient who scored above the 75% percentile should be audited to see why they were worse than 75% of the other patients at that time-point. If words are preferred to numbers, the attributes poor (>90 percentile), below average ( $75 \le 90$ ), average ( $25 \le 75$ ), above average ( $10 \le 25$ ) and excellent ( $\le 10$ ) could be considered for the results of treatment.

Our study has analysed 6336 and 4687 responses to the Oxford hip and knee questionnaires, respectively, over a period of 54 months. A standard for surgical audit has been proposed based on a sample from one hospital and various grades of surgeons. The setting of these standards will need to be reviewed as the sensitivity and specificity of samples from other hospitals and surgeons are published. However, our results give a baseline for comparisons, with the possibility of deviations from the standard forming the basis for audit and survival analysis in the future.

The authors are grateful to the ROH Orthopaedic Charity for helping to fund clerical staff for this project.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

#### References

- Harris WH. Traumatic arthritis of the hip after dislocation in acetabular fractures: treatment by mould arthroplasty. J Bone Joint Surg [Am] 1969;51-A:737-55.
- Charnley J. Long term results of low friction arthroplasty of the hip performed as a primary intervention. J Bone Joint Surg [Br] 1972;54-B:61-76.
- Drake BG, Callahan CM, Dittus RS, Wright JG. Global rating systems in assessing knee arthroplasty outcomes. J Arthroplasty 1994;9:409-17.
- 4. Davies AP. Rating systems for total knee replacement. Knee 2002;9:261-6
- 5. Pynsent PB. Choosing an outcome measure. J Bone Joint Surg [Br] 2001;83-B:792-4.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg [Br] 1996;78-B:185-90.
- Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perception of patients about total knee replacement. J Bone Joint Surg [Br] 1998;80-B:63-9.
- 8. Fairbank JCT, Pynsent PB. The Oswestry disability index. Spine 2000;25:2940-53.
- Ihaka R, Gentleman R. A language for data analysis and graphics. J Comput Graph Stat 1996;5-3:299-314.
- Hanna AW, Pynsent PB, Learmonth DJ, Tubbs ON. A new computer-based triage tool for patients with knee disorders. *Knee* 2000;7:79-86.
- Pynsent PB, Fairbank JCT. Computer interview system for patients with back pain. J Biomed Eng 1989;11:25-9.
- Aderinto J, Cobb AG. Lateral release for patellofemoral arthritis. Arthroscopy 2002; 18:399-403.
- Gidwani S, Tauro B, Whitehouse S, Newman JH. Do patients have to earn total knee arthroplasty? J Arthroplasty 2003;18:199-203.
- 14. 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.
- 15. Davies GM, Watson DJ, Bellamy N. Comparison of the responsiveness and relative effect size of the Western Ontario and McMaster Universities Osteoarthritis Index and the short-form medical outcomes study survey in a randomized, clinical trial of osteoarthritis patients. *Arthritis Care Res* 1999;12:172-9.
- 16. Ehrich EW, Davies GM, Watson DJ, et al. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000; 27:2635-41.
- Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty: results of a survey of 3600 patients from the Swedish Knee Arthroplasty Register. J Bone Joint Surg [Br] 2001;83-B:339-44.
- Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item knee score for use in Sweden. Acta Orthop Scand 2000;71:268-74.
- Fairbank JCT, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271-3.
- Daniel J, Pynsent PB, McMinn D. Survival analysis of metal-on-metal hip resurfacing in young patients with osteoarthritis. J Bone Joint Surg [Br] 2004;86-B:177-84.
- McGrory BJ, Harris WH. Can the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index be used to evaluate different hip joints in the same patient? J Arthroplasty 1996;11:841-4.
- 22. McMurray R, Heaton J, Sloper P, Nettleton S. Measurement of patient perceptions of pain and disability in relation to total hip replacement: the place of the Oxford hip score in mixed methods. *Qual Health Care* 1999;8:228-33.
- Harcourt WG, White SH, Jones P. Specificity of the Oxford knee status questionnaire: the effect of disease of the hip or lumbar spine on patients' perception of knee disability. J Bone Joint Surg [Br] 2001;83-B:345-7.
- Robertsson O, Dunbar MJ. Patient satisfaction compared with general health and disease-specific questionnaires in knee arthroplasty patients. J Arthroplasty 2001;16: 476-82.
- 25. Fitzpatrick R, Morris R, Hajat S, et al. The value of short and simple measures to assess outcomes for patients of total hip replacement surgery. *Qual Health Care* 2000;9:146-50.
- 26. Dawson J, Jameson-Shortall E, Emerton M, et al. Issues relating to long-term follow-up in hip arthroplasty surgery: a review of 598 cases at 7 years comparing 2 prostheses using revision rates, survival analysis, and patient-based measures. J Arthroplasty 2000;15:710-17.
- Palmer SH, Servant CT, Maguire J, Parish EN, Cross MJ. Ability to kneel after total knee replacement. J Bone Joint Surg [Br] 2002;84-B:220-2.
- 28. Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of 'noise' in monitoring patient-based outcomes: generic, disease-specific and site-specific instruments for total hip replacement. J Health Ser Res Policy 1996;1:224-31.