

Quality of life after knee arthroplasty

A randomized study of 3 designs in 42 patients, compared after 4 years

Maria B Hilding^{1,3}, Bengt Bäckbro² and Leif Ryd³

We assessed different yardsticks for outcome 4 (3–5) years after surgery in a prospective, randomized study of 42 patients, where 3 designs of cementless knee prostheses were used. The prognosis with regard to loosening, previously obtained by radiostereometry after 2 years of follow-up, was utilized. Patients with a prognosis of stable implant fixation (two thirds) were compared with those where loosening was predicted (one third). Hospital for Special Surgery score and Visual Analogue Scales regarding pain at rest, "first step" pain, pain during activity and global function, showed consistent postoperative improvements, but no differences between the design and prognosis groups were found. Radiolucent

lines were registered both as yes/no and number of zones. Lines and prognosis were associated, but not lines and design groups. Quality of life assessment by the Nottingham Health Profile questionnaire showed that the poor prognosis group had increased pain and significant disturbances of sleep and emotions, as well as difficulty in enjoying hobbies and holiday activities. No differences were found between the design groups. Altogether, the patients showed profiles comparable to a healthy reference group. We conclude that the Nottingham Health Profile is a sensitive, relevant and simple measure of outcome after knee arthroplasty.

Departments of ¹Orthopedics and ²Radiology, Central Hospital, S-721 89 Västerås. Tel +46 21-17 30 00. Fax -17 52 90; and ³Department of Orthopedics, University Hospital, Lund, Sweden
Submitted 96-04-16. Accepted 97-01-25

Evaluation of results after knee arthroplasty is difficult; failure is largely a matter of definition. Revision is a convenient definition due to the exact occurrence in time. However, prosthetic loosening and other mechanisms of failure do not occur instantaneously, but develop slowly. Criteria for failure in the form of deviations from the desired alignment and prosthetic positions are uncertain, unless pronounced (Ryd 1986, Albrektsson and Herberts 1988). Radiological lucent lines are unreliable, if not wide and complete (Ryd 1986, Ecker et al. 1987, Albrektsson and Herberts 1988, Ritter et al. 1994). Precise radiography, such as radiostereometric analysis (RSA), is a sensitive tool by which loosening can be predicted during the first 1–2 years, before clinical symptoms appear (Freeman and Plante-Bordeneuve 1994, Kärrholm et al. 1994, Ryd et al. 1995, Krismer et al. 1995, Walker et al. 1995).

However, simpler methods are attractive as being more suitable for larger patient groups, if reproducible and provided that their sensitivity and specificity are related to clinical signs of early failure.

Many clinical scoring systems exist, with varying emphasis on pain and function (Drake et al. 1994). They cannot be directly compared, and the capacity for discrimination is generally poor, rendering out-

come evaluations difficult (Andersson 1972, Galante 1985, Insall et al. 1989, Callaghan et al. 1990, Ryd et al. 1997). Quality of life assessments, like the Nottingham Health Profile (Hunt et al. 1980, Hunt and Wiklund 1987), have emerged as a new mode for quantifying health, easy to interpret also for the layman. Patient-based questionnaires have the advantage of focusing on the patient's experienced health problems, which is the more relevant issue in outcome studies (Gartland 1988, Wiklund and Romanus 1991, Rissanen 1995).

We investigated whether simpler methods, such as clinical scores, visual analogue scales of pain, radiographic lucent lines or quantitation of quality of life, could detect tibial components at risk of loosening.

Patients and methods

45 patients with gonarthrosis were entered consecutively in a prospective, randomized study at the Department of Orthopedics, Central Hospital of Västerås, during the period 1989–1992. This patient series has been reported earlier (Hilding et al. 1993, 1995). The study was approved by the Ethics Committee of the Karolinska Institute, Stockholm, and the

patients gave their informed consent. The following criteria were used: age 60–75 years, arthrosis stages III–V according to Ahlbäck (1968), no previous fracture, osteotomy or arthroplasty of the investigated knee, no steroid medication, no systemic illness affecting the skeleton, no joint complaints apart from the knees. Bilateral involvement was accepted due to the low incidence of severe, monarticular gonarthrosis, but only one knee from each patient was operated on in the study. However, 7 of the patients had been fitted with different kinds of knee prostheses before this study and 7 were operated on with TKA 6 months–2 years after the study began. The rest had either unilateral gonarthrosis or insufficient symptoms to require an arthroplasty. The patients were randomized by use of closed envelopes, in groups of 3, to 3 different designs of uncemented tricompartmental prostheses: the Tricon-M and Tricon stem (Smith & Nephew, Memphis, TN) and the PCA resurfacing (Howmedica, Rutherford, NJ). The Tricon prostheses were identical, except for the presence or not of a central stem on the tibial component. Patella components were used in all cases. Two experienced surgeons performed the operations by a standardized technique. Postoperative complications were minor, 1 deep venous thrombosis and 1 transient confusion. No infections occurred. 2 patella components, 1 from each of the Tricon groups, were removed due to plastic wear after 2 years, with relief of symptoms.

All evaluations before and after the operations were performed by an independent observer (the first author). 3 patients, who died of unrelated causes 6 months, 1 and 2 years after surgery, were excluded. Otherwise, all patients attended all scheduled check-ups. Clinical scores were obtained using the Hospital for Special Surgery Score (HSS) (Insall et al. 1976), preoperatively and at 6 months, 1, 2, and 4 years (range 3–5 years) postoperatively. At the same time intervals, 4 modes of pain (pain at rest, first-step pain, pain during activity and global satisfaction with the investigated knee) were recorded by the patients following standardized instructions, on Visual Analogue Scales (VAS) (Huskisson 1974). Activity level was graded from 1–10 on a scale developed for hip replacements (Amstutz et al. 1984). Activity levels were recorded postoperatively at 1, 2, and 4 (3–5) years. At the latest follow-up, 4 (3–5) years postoperatively, the Nottingham Health Profile (NHP) (Hunt et al. 1980, Hunt and Wiklund 1987) quality of life questionnaire was used (n 41). The NHP questionnaire consists of 2 parts. The first 38 yes/no statements are about perceived health problems in the categories of pain, sleep, energy, physical mobility, social isolation and emotional reactions. The second

part contains 7 single yes/no statements whether daily function is affected by health problems, in home tasks, paid employment, social life, family relationships, sex life, hobbies, or recreational and holiday activities. These categories are not combined in a single overall estimate of general health, but represent the various dimensions of life separately. The English version has been tested extensively for reliability and validity (Hunt et al. 1980). A Swedish translation has been available since 1987, with a high degree of congruency in test results (Hunt and Wiklund 1987).

Radiographic data were collected, such as leg alignment, Hip-Knee-Ankle angle (HKA), prosthetic component positions, the coverage of the tibial component of the cut surface, off-set in placement and joint line position (Hilding et al. 1993). The radiolucent lines below the tibial baseplate were recorded yearly after surgery, with 0.5 mm increments and in separate zones, distributed according to recommendations of the Knee Society (Ewald 1989), modified so that stems formed single zones.

RSA was also performed on all patients and the results of 2 years of follow-up have been published (Hilding et al. 1995).

RSA has been reported to have a predictive value of 80% in detecting tibial components at risk of loosening after only 2 years of follow-up (Ryd et al. 1995). This capacity was utilized for prognostication, by using 200 μ m migration in the second postoperative year as a threshold. The 42 patients still alive 3 years after surgery were thus divided into 2 groups: 27 with good prognosis concerning fixation of the tibial component and 15 at risk of loosening (Hilding et al. 1995), regardless of prosthesis design. In addition, the 3 groups of patients with different prosthesis designs were compared.

As previously reported (Hilding et al. 1995), age, stage of arthrosis, sex, weight and alignment, pre- or postoperative, were not significantly different between the groups with various prognoses. The prosthesis design groups showed no differences preoperatively or in demographics, but a mean difference of 2° was found in postoperative alignment between the PCA and the Tricon groups.

The statistical analysis was done with non-parametric methods, the Mann-Whitney U-test for comparisons of 2 independent groups and the Kruskal-Wallis test for comparisons of 3 independent groups. Comparisons of values from the same group on different measurement occasions were done with the Wilcoxon signed ranks test. Categorical data were analyzed by the chi-square test. P-values < 0.05 were considered significant.

Table 1. The Nottingham Health Profile in this study at 4 years postoperatively and in a previous study

	n	NHP 1							NHP 2				
		Pain	Energy	Sleep	Mobility	Emotions	Social isolation	House work	Holidays	Hobbies	Social life	Sex life	Family life
<i>This study</i>													
all patients	41	14	15	17	16	8	5	27	22	24	7	12	2
good prognosis	26	8	12	10	14	6	5	19	8	8	4	8	0
poor prognosis	15	24	20	29	18	12	5	40	40	47	13	20	7
<i>Wiklund and Romanus 1991</i>													
healthy reference		8	15	16	6	9	5	18	14	15	14	14	8
THR preop.	56	75	63	55	48	30	14	77	71	64	46	34	23
THR 1 yr postop.	56	15	20	19	18	10	7	30	25	23	7	9	2

THR total hip replacement.

The 45 Nottingham Health Profile questions can only be answered by yes or no. Part 1 comprises 38 questions, where "no" answers generate zero points and "yes" answers are weighted according to their relative contribution to each of the 6 different profiles. Each profile in part 1 may generate 0–100 points. Part 2 consists of 7 statements, whether certain areas of life are affected by health. (Paid employment not used in this study.) An affirmative answer scores one, a negative zero. The sum in the cell is made up of the percentage of "yes" answers.

Results

No significant differences were found in postoperative activity levels between the 3 prosthesis groups. The median level at 2 years after surgery was 6 in both the Tricon stem and the PCA groups and 7 in the Tricon-M group. Approximately 2 years later (3–5 years after surgery), no significant changes had occurred, the median level was 6 in all groups. Patients in the groups with higher activity levels did not show significantly more migration in the second year after surgery, as measured by RSA, than patients with a more sedentary life-style. Nor was there any difference in activity level between the good and the poor prognosis groups.

HSS clinical scores improved from a median of 60 preoperatively to 82 6 months after surgery ($p < 0.001$). With time, some further improvements occurred, median 86 at 1 year ($p < 0.001$), 89 at 2 years ($p = 0.006$) and 88 at 4 (3–5) years after surgery (NS). There were no significant differences between the prosthesis groups, except at 4 (3–5) years, when the Tricon stem group had deteriorated slightly to a median of 82 points compared to 87.5 for Tricon-M and 90 for PCA ($p = 0.02$). When the prognosis of fixation based on RSA was considered and the material was divided into a good (n 27) and a poor (n 15) prognosis group, the HSS scores showed no significant differences between these groups.

The 4 visual analogue scales showed significant reductions in pain after the operation, most of the effect being present at 6 months, but some further reductions were noted up to 1 or 2 years. Thereafter, no significant changes occurred. The 3 prosthesis designs showed no differences for pain at rest, "first-step

pain" and global satisfaction with the investigated knee, but for pain during activity at 4 (3–5) years, the Tricon stem group showed a median value of 5 mm (VAS scale 0–100 mm), both other groups 0 mm ($p = 0.04$). The 2 groups with different prognoses showed no differences in any of the 4 scales, at any time.

The Nottingham Health Profile in the whole group of patients 4 (3–5) years after surgery showed answers in part 1 (sleep, energy, social isolation and emotional reactions), similar to those of a healthy population (Wiklund and Romanus 1991), except for pain and mobility profiles, where our patients showed slightly increased levels (Table 1). In part 2, home tasks, hobbies, recreational activities or holidays, were slightly more affected than in the healthy reference population (Wiklund and Romanus 1991), contrary to social life, family relationships and sex life, which were less affected (Table 1). Paid employment was not applicable, since all patients were retired. The 3 prosthesis groups could not be distinguished by the Nottingham Health Profile ($p = 0.2–0.9$). Significant differences were found when the prognosis regarding future loosening was considered. In part 1, the pain profile showed the largest difference ($p = 0.002$), but also sleep ($p = 0.04$) and emotions ($p = 0.04$) were more affected in the risk group (Table 1). In part 2, the risk group had more difficulty in enjoying hobbies ($p = 0.004$) and holiday activities ($p = 0.03$) (Table 1).

The radiographic analysis showed that 26 patients, consistently during the follow-up period showed no lucent lines at all or less than 0.5 mm. Radiolucent lines once established generally remained, but in 3 cases (2 PCA cases, around the pegs, and 1 Tricon stem), the width of incomplete lucent lines was reduced with time. The 16 patients with lucent lines in 1

Table 2. Radiographic lucent lines according to prognosis group

Prognosis	Lucent lines	
	yes	no
Poor	9	6
Good	7	20

Lucent lines, yes/no, refers to the presence of any radiographic lucent lines at 4 years, > 0.5 mm, disregarding the number of zones involved. The prognosis, good/poor, refers to the level of migration the second year, as measured by RSA.

or more zones wider than 0.5 mm did not show significantly more migration in the second year, as measured by RSA (mean 295 μ m), compared to the 26 patients with thinner or no lines (mean 175 μ m) ($p = 0.3$). The presence or not of radiolucent lines > 0.5 mm was related to the prognosis ($p = 0.03$) (Table 2). The number of zones with lucent lines > 0.5 mm at 4 (3–5) years and the prognosis group were also related ($p = 0.03$).

Discussion

No patient in this study was classified as having a poor result (according to HSS levels), nor did anyone report dissatisfaction with the operation. A follow-up period of about 4 years is short for conclusions about the performance of knee replacements, since normally few failures occur during that time. When blunt tools like scoring systems are used, only very unsatisfactory results can be detected early. The assessment of arthroplasties by composite scores has been challenged (Galante 1985, Gartland 1988, Callaghan et al. 1990, Drake et al. 1994, Nilsson et al. 1994). Most scores lack validation and include various parameters which are weighted differently (Callaghan et al. 1990, Drake et al. 1994). The reproducibility of the individual measurements is often poor (Ryd et al. 1997) and aggravates the overall uncertainty. Improvement of the methods used to evaluate outcome could be very important, both to permit comparisons and reduce bias, but also to detect subtle early symptoms. Major advantages of the NHP registration are the thorough validations performed and the elimination of bias by the examiner, as well as the focus on quality of life.

NHP results for all patients in this study were similar to the profiles in 2 studies on successful hip replacements (Wiklund and Romanus 1991, Nilsson et al. 1994), slightly better than the results from a study 2 and 5 years after knee replacement (Rissanen et al. 1995), and comparable to a healthy population (Wik-

lund and Romanus 1991). Among 49 total hip replacements reported by Nilsson et al. (1994), 8 patients had radiological signs of femoral component loosening 5 years after surgery, and this subgroup showed significant reductions in function and quality of life, as measured by NHP. The signs and symptoms of loosening were not severe enough to recommend revision and the Charnley hip score was the same as in the rest of the patients. Our study shows analogous results. The group predicted as future failures by RSA scored no worse than the others, nor did they have increased pain levels using VAS, but they differed several of the dimensions of the NHP. The discrepancy was greatest in the pain profile, but also sleep and emotions were significantly more affected in the risk group, together with difficulties in the practice of hobbies and holiday activities. We interpret these differences as a reflection of early clinical symptoms, not detected by HSS or VAS.

VAS scales in the measurement of pain have been reported to be robust, reliable and sensitive (Huskisson 1974, Revill et al. 1976, Scott and Huskisson 1976). Despite thorough explanations of VAS, in the same manner and by the same observer at each examination, we felt that some patients had difficulties in understanding how to answer. The patients may better express such minor and easily overlooked symptoms when distinct questions are posed, concerning situations in everyday activities and feelings. The dichotomous nature of the NHP questions may force the patient into a definite choice: yes or no. However, NHP includes symptoms not only from the involved knee, in contrast to HSS and VAS.

Radiolucent lines are often reported as signs of loosening and a relationship has sometimes been shown in the literature (Ecker et al. 1987, Ewald 1989, Ritter et al. 1994). This study does not conflict with these findings, indeed a relationship was found. However, lucent lines are easily missed because they depend on the quality of the radiographs, subject to observer bias and are difficult to classify. Conclusions from thin lucent lines in the individual case are therefore difficult to draw. Nevertheless, due to the absence of more reliable tools, such assessments are widely made.

Acknowledgements

The authors thank Dr. Stefan Sörensen, Department of Research, IAMU, University of Uppsala, Central Hospital, Västerås, for statistical advice. Financial support was supplied by the County Council of Västmanland Research Fund, the Odd Fellow Foundation, Greta and Johan Kock's Stiftelse, Stiftelsen Bistånd åt Vanföra i Skåne, Konung Gustav

V's 80-årsfond, Alfred Österlunds Stiftelse, The Medical Faculty of Lund University and the Medical Research Council Project 09509.

References

- Ahlbäck S. Osteoarthritis of the knee. A radiographic investigation. *Acta Radiol Scand (Suppl 277)* 1968; 1-71.
- Albrektsson B E J, Herberts P. ICLH knee arthroplasty. A consecutive study of 108 knees with uncemented tibial component fixation. *J Arthroplasty* 1988; 3 (2): 145-56.
- Amstutz H C, Thomas B J, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg (Am)* 1984; 66 (2): 228-41.
- Andersson G. Hip assessment: A comparison of nine different methods. *J Bone Joint Surg (Br)* 1972; 54 (4): 621-5.
- Callaghan J J, Dysart S H, Savory C F, Hopkinson W J. Assessing the results of hip replacement. A comparison of five different rating systems. *J Bone Joint Surg (Br)* 1990; 72 (6): 1008-9.
- Drake B G, Callahan C M, Dittus R S, Wright J G. Global rating systems used in assessing knee arthroplasty outcomes. *J Arthroplasty* 1994; 9 (4): 409-17.
- Ecker M L, Lotke P A, Windsor R E, Cella J P. Long-term results after total condylar knee arthroplasty. Significance of radiolucent lines. *Clin Orthop* 1987; 216: 151-8.
- Ewald F C. The Knee Society total knee arthroplasty roentgenographic evaluation and scoring system. *Clin Orthop* 1989; 248: 9-12.
- Freeman M A R, Plante-Bordeneuve P. Early migration and late aseptic failure of proximal femoral prostheses. *J Bone Joint Surg (Br)* 1994; 76 (3): 432-8.
- Galante J. Editorial. The need for a standardized system for evaluating results of total hip surgery. *J Bone Joint Surg (Am)* 1985; 67 (4): 1357-64.
- Gartland J J. Orthopaedic clinical research. Deficiencies in experimental design and determinations of outcome. *J Bone Joint Surg (Am)* 1988; 70 (9): 1357-64.
- Hilding M B, Asplund S, Bäckbro B O, Ryd L. Alignment accuracy by two different instrumentation systems in total knee arthroplasty. A prospective randomized study of 45 patients. *Am J Knee Surg* 1993; 6 (4): 150-66.
- Hilding M B, Yuan X, Ryd L. The stability of three different cementless tibial components. A randomized radiostereometric study in 45 knee arthroplasty patients. *Acta Orthop Scand* 1995; 66 (1): 21-7.
- Hunt S M, Wiklund I. Cross-cultural variation in the weighting of health statements: a comparison of English and Swedish valuations. *Health Policy* 1987; 8: 227-35.
- Hunt S M, McKenna S P, McEwen J, Backett E M, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. *J Epidemiol Community Health* 1980; 34: 281-6.
- Huskisson E C. Measurement of pain. *Lancet* 1974; 2: 1127-31.
- Insall J N, Ranawat C S, Aglietti P, Shine J. A comparison of four models of total knee replacement prostheses. *J Bone Joint Surg (Am)* 1976; 58 (6): 754-65.
- Insall J N, Dorr L D, Scott R D, Scott W N. Rationale of the Knee Society clinical rating system. *Clin Orthop* 1989; 248: 13-4.
- Krismer M, Bauer R, Fischer M. Prediction of the final outcome of hip arthroplasty by measurement of migration. *J Bone Joint Surg (Br)* (Suppl II) 1995; 77: 163.
- Kärholm J, Borssén B, Löwenhielm B, Snorrason F. Does early micromotion of femoral stem prostheses matter? 4-7 year stereoradiographic follow-up of 84 cemented prostheses. *J Bone Joint Surg (Br)* 1994; 76 (6): 912-7.
- Nilsson L T, Franzén H, Carlsson Å S, Önnertfalt R. Early radiographic loosening impairs the function of a total hip replacement. The Nottingham Health Profile of 49 patients at five years. *J Bone Joint Surg (Br)* 1994; 76 (2): 235-9.
- Revill S I, Robinson J O, Rosen M, Hogg M I J. The reliability of a linear analogue for evaluating pain. *Anaesthesia* 1976; 31: 1191-8.
- Rissanen P, Aro S, Slätis P, Sintonen H, Paavolainen P. Health and quality of life before and after hip or knee arthroplasty. *J Arthroplasty* 1995; 10 (2): 169-75.
- Ritter M A, Herbst S A, Keating E M, Faris P M. Radiolucency at the bone-cement interface in total knee replacement. The effects of bone-surface preparation and cement technique. *J Bone Joint Surg (Am)* 1994; 76 (1): 60-5.
- Ryd L. Micromotion in knee arthroplasty. A roentgen stereophotogrammetric analysis of tibial component fixation. *Acta Orthop Scand (Suppl 220)* 1986; 57: 1-80.
- Ryd L, Albrektsson B E J, Carlsson L, Herberts P, Lindstrand A, Regné L, Dansgaard F, Toksvig-Larsen S. Roentgen stereophotogrammetric analysis (RSA) as a predictor of mechanical loosening. *J Bone Joint Surg (Br)* 1995; 77 (3): 377-83.
- Ryd L, Kärholm J, Ahlvin P, The Score Assessment Group. Knee scoring systems in gonarthrosis. Evaluation of inter-observer variability and the envelope of bias. *Acta Orthop Scand* 1997; 68 (1): 41-5.
- Scott J, Huskisson E C. Graphic representation of pain. *Pain* 1976; 2: 175-84.
- Walker P S, Mai S F, Cobb A G, Bentley G, Hua J. Prediction of clinical outcome of THR from migration measurements on standard radiographs. A study of cemented Charnley and Stanmore femoral stems. *J Bone Joint Surg (Br)* 1995; 77 (5): 705-14.
- Wiklund I, Romanus B. A comparison of quality of life before and after arthroplasty in patients who had arthrosis of the hip joint. *J Bone Joint Surg (Am)* 1991; 73 (5): 765-9.