

No difference in health-related quality of life in hip osteoarthritis compared to degenerative lumbar instability at pre- and 1-year postoperatively

A prospective study of 101 patients

Ole Juul¹, Freyr G Sigmundsson¹, Ole Ovesen¹, Mikkel O Andersen¹, Carsten Ernst² and Karsten Thomsen¹

Departments of Orthopedics, ¹University Hospital of Odense, Odense, ²Esbjerg Hospital, Esbjerg, Denmark
Correspondence KT: Karsten.Thomsen@ouh.fyns-amt.dk
Submitted 05-03-13. Accepted 06-03-29

Background Total hip replacement (THR) is a very successful and refined surgical procedure when compared to crude bony fusion in degenerative lumbar segmental instability (LF). We compared the pre- and postoperative health-related quality of life status of THR and LF patients.

Patients and methods We prospectively studied 51 THR patients and 50 LF patients. The outcome parameters were SF-36 and Oswestry Disability Index (ODI), measured preoperatively and at 1 year postoperatively. The status of the patients was compared to that of an age-matched healthy control group.

Results The preoperative SF-36 and ODI scores were similar between the groups, except for the subscale role emotional. One year postoperatively, only the differences in 3 subscales (physical functioning, role physical, and role emotional) and in the standardized physical component reached statistical significance; the THR-patients scored worse than the LF-patients. The improvements in SF-36 and ODI reached statistical significance in both groups.

Interpretation The differences in quality of life between the THR and LF patients were similar pre- and postoperatively. The quality of life of both cohorts improved considerably and significantly after the treatment, but they remained at a level significantly below that of a general age-matched population.

One of the factors thought to be responsible for the success of total hip replacement (THR) is the fact that there are generally well-defined indications for surgery. In Denmark, the annual incidence of THR per 100,000 inhabitants was 93 in 1998 and the mean age at the primary THR was 68 years (Lucht 2000). About 75% of these patients had a primary THR because of primary osteoarthritis, and this disease is again well-defined, the average patient having little functional characteristics or non-organic complaints in his/her clinical appearance (Arokoski et al. 2004).

The incidence of surgical fusion of painful segments in the degenerated lumbar spine (LF) has increased almost exponentially in many countries of the western world. In the United States, the incidence is currently 50 LF per 100,000 inhabitants per year (Lipson 2004). The rationale behind the fusion in degenerative lumbar segmental instability is to eliminate pathological motion in order to relieve pain. In contrast to THR, LF is the subject of controversy both within and outside the field of orthopedic surgery (Deyo et al. 2004). In spite of a major research effort, many aspects of this treatment are thought to be less evidence-based than in THR (Nachemson et al. 1996, Nachemson 1999, Hägg et al. 2002). LF presents high practice variation both at the national and international levels (Cherkin et al. 1994). One concern is that the outcome in general will be poorer the higher the rate

of LF, since this has been shown to be the case for lumbar disc herniation and spinal stenosis surgery (Keller 1999). Another important point of concern is the indication for surgery, because many of these patients suffer significant psychosocial disturbances (Mendelson 1984, Nachemson 1999, Rush et al. 2000).

We compared the pre- and postoperative health-related quality of life status in two cohorts of patients, the first having a THR for primary osteoarthritis (THR patients) and the other a lumbar spinal fusion for degenerative segmental instability (LF patients) at the same department of orthopedic surgery. The patients' status was compared to age-matched healthy controls.

Patients and methods

THR patients

54 THR patients suffering from primary osteoarthritis were included prospectively. 1 patient withdrew his consent at the 1-year follow-up, 1 patient was excluded due to 6 logical errors in the SF-36 score, and 1 patient died in the postoperative period from pulmonary embolism. This left 51 patients (28 men) with a mean age of 65 (56–87) years in this group.

Surgical technique

All patients were operated with a posterolateral incision. As a rule, patients older than 65 years had a cemented THR, patients aged between 60 and 65 years had a hybrid THR, and patients younger than 55 years had an uncemented THR. All patients received prophylactic dicloxacillin (2 g, i.v.) intraoperatively, and 1 g i.v. 8, 16, and 24 h later. All patients received LMW-heparin (enoxaparin 40 mg once a day) until discharge.

LF patients

From the 54 prospectively included LF patients, 3 withdrew their consent at the follow-up and 1 patient had numerous unanswered questions on the preoperative SF-36 questionnaire, so the study group consisted of 50 patients (26 women), mean age 53 (20–70) years. All suffered lumbar or lumbosacral degenerative instability, defined by severe, chronic motion-induced back pain due to

primary degeneration or degeneration after prior decompressive surgery and resistant to conservative treatment for a minimum of 6 months. The indication for fusion was decided before surgery through repeated clinical and neurological examinations, a plain radiograph, and an MRI study. The criteria for segmental degenerative instability on MRI were primarily reduced height of the discus, intervertebral translatory displacement, and Modic signs types I or II.

All patients underwent a posterior/anterior procedure using a rigid pedicle screw fixation system to support the posterolateral fusion and a Brantigan cage as the anterior intracorporeal support. We used autologous bone transplantation from the iliac wing, and hemilaminectomy for neural decompression when indicated, preserving as much of the stabilizing posterior structures as possible (Christensen et al. 2002). All patients received prophylactic dicloxacillin 1 g i.v. intraoperatively, and this was repeated 3 times a day for 5 days. All patients received LMW-heparin (enoxaparin 40 mg once a day) until discharge.

Exclusion criteria for both groups of patients were age less than 20 years, metabolic bone disease, and for the LF patients, previous fusion. No patients were excluded because of co-morbidity or psychosocial instability.

Perceived health status: SF-36

The Medical Outcome Study Short Form 36 (SF-36) was constructed for group comparisons involving generic health concepts not specific to any age, disease or treatment group (Ware and Sherbourne 1992). The SF-36 measures both physical and mental health components over the previous week, covering 8 dimensions (subscales): Physical Functioning (PF), Social Functioning (SF), Role Physical (RP), Role Emotional (RE), Mental Health (MH), Vitality (VT), Bodily Pain (BP), and General Health. For each dimension, item scores are coded, summed, and transformed to a scale from 0 (worst health, maximum disability) to 100 (best health, no disability). The SF-36 establishes two summary measures: the Standardized Physical Component scale (PCS) and the Standardized Mental Component Scale (MCS). A computer was used to calculate the scores for the 8 subscales according to the algorithm described by the SF-36 manual.

Table 1. SF-36 and ODI scale scores for the THR and LF patients, preoperatively and at 1-year postoperatively

Scale	Preoperatively		P ^a	Postoperatively		P ^b	P ^c	P ^d
	THR patients	LF patients		THR patients	LF patients			
SF-36								
Physical Functioning (PF)	36 (32–40)	40 (35–45)	0.2	44 (40–48)	52 (48–57)	0.01	<0.001	0.04
Social Functioning (SF)	53 (47–59)	52 (46–58)	0.8	56 (53–60)	56 (53–60)	0.9	0.01	<0.001
Role Physical (RP)	7.4 (1–14)	17 (9–25)	0.05	33 (26–40)	56 (48–63)	<0.001	<0.001	0.04
Role Emotional (RE)	32 (25–39)	47 (34–60)	0.05	54 (45–63)	75 (66–83)	<0.001	<0.001	<0.001
Mental Health (MH)	67 (62–73)	72 (67–77)	0.3	74 (70–78)	76 (72–80)	0.5	<0.001	<0.001
Vitality (VT)	51 (46–57)	50 (43–58)	0.8	57 (53–60)	58 (53–63)	0.7	0.002	<0.001
Bodily Pain (BP)	36 (31–42)	35 (27–43)	0.7	52 (48–56)	56 (50–62)	0.3	<0.001	<0.001
General Health (GH)	65 (60–71)	66 (62–71)	0.8	73 (69–77)	69 (65–73)	0.2	<0.001	<0.001
Standardized Physical Component (PCS)	31 (30–33)	31 (29–33)	0.9	36 (35–38)	39 (37–40)	0.03	<0.001	<0.001
Standardized Mental Component (MCS)	46 (44–48)	48 (46–50)	0.1	49 (47–51)	51 (49–52)	0.2	<0.001	<0.001
Oswestry Disability Index (ODI)	30 (0.7)	31 (0.7)	0.2	15 (0.6)	14 (0.7)	0.2	0.02	<0.001

Data are mean (95% CI)

^a p-values of differences between the THR- and the LF-patients preoperatively.

^b p-values of the differences between the THR- and the LF-patients postoperatively.

^c p-values of the differences from pre- to 1-year postoperatively in the THR-patients.

^d p-values of the differences from pre- to 1-year postoperatively in the LF-patients.

Oswestry Disability Index (ODI)

The ODI (Version 2.0) is a 10-question, low back-specific instrument designed to measure disability in spine patients (Fairbank et al. 1980). In the ODI, 0 indicates no disability and 500 represents maximum disability, and the figures were converted to percentage of disability according to Fairbank et al. (1980). The SF-36 and the ODI have both been validated recently for detection of changes in chronic pain treatment not confined to the spine, and both showed good psychometric properties (Witting et al. 2004).

Statistics

The significance of changes within groups before and after operation was determined by the paired Student's t-test, and the significance of differences in values between groups by Student's t-test for 2 independent samples, both using SPSS software version 612.0. The results were considered significant at $p < 0.05$.

Results

For the baseline preoperative scores neither the 8 subscales, nor the SF-36 summary measures, nor the ODI showed any significant difference between the groups (Table 1). At 1 year postoperatively, only the differences in the 3 subscales PF, RP, and RE, and in the standardized physical component, PCS, reached statistical significance. For all of these, the THR patients scored worse than the LF patients. For all parameters, the improvements in SF-36 and ODI from before surgery to the 1-year follow-up in both the THR patients and the LF patients reached statistical significance.

When comparing both the THR and the LF cohorts, both preoperatively and postoperatively, to two age-matched Danish control cohorts (Tables 2 and 3) (Bjørner et al. 1997), we found significantly lower SF-36 scores in the patients for all variables except for the subscale GH in the THR patients preoperatively. The SF-36 status of the patients relative to the Danish norm is given in Table 4.

Table 2. SF-36 and ODI scale scores for the THR patients and Danish controls

Scale	THR patients preoperatively	THR patients postoperatively	Danish sample 65–74 years ^a	P-value ^b	P-value ^c
SF-36					
Physical Functioning (PF)	36 (32–40)	44 (40–48)	74 (71–76)	<0.001	<0.001
Social Functioning (SF)	53 (47–59)	56 (53–60)	90 (88–92)	<0.001	<0.001
Role Physical (RP)	7.4 (1–14)	33 (26–40)	63 (59–67)	<0.001	<0.001
Role Emotional (RE)	32 (25–39)	54 (45–63)	78 (75–82)	<0.001	<0.001
Mental Health (MH)	67 (62–73)	74 (70–78)	82 (81–84)	<0.001	0.002
Vitality (VT)	51 (46–57)	57 (53–60)	69 (67–71)	<0.001	<0.001
Bodily Pain (BP)	37 (31–42)	52 (48–56)	78 (76–80)	<0.001	<0.001
General Health (GH)	65 (60–71)	73 (68–76)	66 (64–68)	0.867	0.036
Standardized Physical Component (PCS)	31 (30–33)	36 (35–38)	45 (44–46)	<0.001	<0.001
Standardized Mental Component (MCS)	46 (44–48)	49 (47–51)	56 (56–57)	<0.001	<0.001
Data are mean (95% CI)					
^a n=427					
^b p-values of the differences between the THR patients preoperatively and the age-matched Danish sample.					
^c p-values of the differences between the THR patients at 1-year postoperatively and the age-matched Danish sample.					

Table 3. SF-36 and ODI scale scores for the LF patients and Danish controls

Scale	LF-patients preoperatively	LF-patients postoperatively	Danish sample 45–54 years ^a	P-value ^b	P-value ^c
SF-36					
Physical Functioning (PF)	40 (35–45)	52 (47–57)	88 (86–89)	<0.001	<0.001
Social Functioning (SF)	52 (46–58)	56 (52–60)	91 (89–92)	<0.001	<0.001
Role Physical (RP)	10 (9–25)	56 (48–63)	83 (81–86)	<0.001	<0.001
Role Emotional (RE)	47 (34–60)	75 (66–83)	87 (85–89)	<0.001	0.002
Mental Health (MH)	72 (67–77)	76 (72–80)	82 (81–84)	<0.001	0.007
Vitality (VT)	50 (43–58)	58 (53–63)	71 (69–72)	<0.001	<0.001
Bodily Pain (BP)	35 (27–43)	56 (50–62)	78 (76–80)	<0.001	<0.001
General Health (GH)	66 (62–71)	69 (65–73)	75 (73–77)	0.004	0.039
Standardized Physical Component (PCS)	31 (29–33)	39 (37–40)	50 (50–51)	<0.001	<0.001
Standardized Mental Component (MCS)	48 (46–50)	51 (49–52)	54 (54–55)	<0.001	0.009
Data are mean (95% CI)					
^a n=700					
^b p-values of the differences between the LF patients preoperatively and the age-matched Danish sample.					
^c p-values of the differences between the LF patients at 1-year postoperatively and the age-matched Danish sample.					

Discussion

During the past decade, interest in assessing how patients perceive the effect of a disease has increased. The so-called health-related quality of life (HRQoL) assessment has become a significant outcome parameter in the medical literature, and SF-36 is widely accepted as a reliable tool in this connection (Leidy et al. 1999).

A recent Danish study on the epidemiology of chronic non-malignant pain used the SF-36 on 12,333 well-defined individuals, either normal controls or patients suffering from chronic pain (Eriksen et al. 2003). They found an overall prevalence of chronic pain of 19%, and compared to the control group the chronic pain group showed consistently lower scores in all subscales of the self-reported SF-36 assessment.

Table 4. SF-36 corrected for age in the two treatment groups

Scale	THR patients preoperatively	LF patients preoperatively	THR patients postoperatively	LF patients postoperatively
SF-36				
Physical Functioning (PF)	0.488	0.453	0.600	0.597
Social Functioning (SF)	0.592	0.573	0.629	0.614
Role Physical (RP)	0.117	0.205	0.520	0.668
Role Emotional (RE)	0.409	0.536	0.685	0.857
Mental Health (MH)	0.819	0.872	0.900	0.922
Vitality (VT)	0.746	0.711	0.820	0.819
Bodily Pain (BP)	0.468	0.447	0.667	0.720
General Health (GH)	0.992	0.885	1.103	0.919
Standardized Physical Component (PCS)	0.687	0.619	0.800	0.768
Standardized Mental Component (MCS)	0.826	0.890	0.883	0.937

Osteoarthritis of the hip or degenerative lumbar segmental instability may both lead to chronic pain; thus, the orthopedic surgeon may offer a THR for the first group and an LF for the latter type of patient. Between 1996 and 2001, the annual number of THRs increased by 14%—as compared to an increase of 77% for LH in the United States, so that in 2001 the total number of operations was 300,000 for each of the two operations (Agency 2004).

THR is generally considered to be a very successful orthopedic surgical procedure. On the other hand, LF is a much more controversial procedure and its usefulness is still questioned. The prejudice amongst orthopedic surgeons and many others dealing with these patients may be that the patients suffering chronic low back pain are victims of more psychosocial co-morbidity than the patients suffering chronic hip pain (Fanuele et al. 2000). It has also been documented that a personality characterized by neuroticism, or having personality disorders, is a significant predictor of poorer functional improvement after LF (Ekselius 1996, Hägg et al. 2003).

We have found no previous prospective studies comparing the perioperative HRQoL in THR and LF. It is important to appreciate that both of the cohorts comprised patients whom the surgeons had selected for surgical treatment. Our study does not predict the extent to which patients as a whole suffering hip pain differ from patients as a whole suffering back pain. The two study cohorts expected to benefit from surgery are both small

parts of the much larger cohorts of hip and back patients. In the LF group, many patients showing personality disorders, widespread pain, negative coping, psychological distress etc. were excluded from surgery. This is no doubt the case also in the THR group, but one could speculate that identification of osteoarthritis of the hip joint is so explicit as compared to degenerative lumbar instability—and the overall results in THR so successful—that there is less emphasis on the negative psychological and sociological predictors in these patients. In our study, both the physical and the psychological profiles were similar in the two treatment groups. Though not statistically significant, the LF patients did in fact score better in the mental health components than did the THR patients whereas the THR patients had the highest scores for the physical health components. These findings should not be overestimated in this relatively small study, but indicate that spinal surgeons have learned how to identify the patients who may benefit most from an LF.

Regarding the outcome, we have been unable to find any studies comparing THR and LF. Hozack et al. (1997) compared the pre- and postoperative SF-36 scores for patients who had THR, total hip revision, total knee replacement, lumbar laminectomy (not LF), and scoliosis surgery. They reported that THR and lumbar laminectomy gave the best results. We found that the THR and LF patients had approximately the same HRQoL one year postoperatively with the LF patients improving moderately more from the treatment than their

THR counterparts. Despite these improvements, both groups of patients had a level of HRQoL that was significantly below the general age-matched population one year postoperatively—with the scores for the physical health components at lower levels than for the mental health components, for both THR and LF.

In a recent study Witting et al. (2004) compared the redundancy, reliability and responsiveness to change of SF-36, ODI, and Multidimensional Pain Inventory (MPI) and concluded that they each had good psychometric properties. They found, however, that SF-36 was superior and that ODI was the least expedient of the 3, and it was concluded that only SF-36 and MDI make unique and complementary contributions to assessment of patients with chronic pain. The ODI data in the present study may therefore not be as valuable as the SF-36 data. We saw a decrease in pain by ODI from an average of 30 preoperatively to an average of 15 postoperatively in both groups. This brings both groups of patients from ‘moderate disability’ to the lowest group ‘minimal disability’ in the 5-group stratification of the ODI by Fairbank et al. (1980).

Contributions of authors

OJ participated in the entire process—planning and execution of the study, data collection, statistical analysis and preparation of the manuscript. FGS assisted in execution of the study, data collection and statistical analysis, and in preparation of the manuscript. OO performed many of the THRs and assisted in conduction of the study, especially in preparation of the manuscript. MOA and CE performed many of the LFs and assisted in conduction of the study, especially in preparation of the manuscript. KT performed many of the LFs and participated in the entire process: planning and execution of the study, data collection, statistical analysis and preparation of the manuscript.

No competing interests declared.

Abraham D J, Herkowitz H N, Katz J N. Indications for thoracic and lumbar spine fusion and trends in use. *Orthop Clin North Am* 1998; 29 (4): 803.

Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project, HCUPnet. (Accessed January 22, 2004, at <http://www.ahrq.gov/data/hcup/>.)

Arokoski M H, Haara M, Helminen H J, Aarakoski J P. Physical function in men with and without hip osteoarthritis. *Acta Phys Med Rehab* 2004; 85 (4): 574-81.

Bjørner J B, Damsgaard M T, Watt T, Bech P, Rasmussen N K, Kristensen T S, Modvig J, Thuneborg K. Dansk manual til SF-36. Lægemedelindustriforeningen, Lif 1997. ISBN: 87-90350-07-3.

Cherkin D C, Deyo R A, Wheeler K, Ciol M A. Physician variation in diagnostic testing for low back pain. Who you see is what you get. *Arthritis Rheum* 1994; 37 (1): 15-22.

Christensen F B, Hansen E S, Eiskjaer S, Hoy K, Helmig P, Neumann P, Bunge C E. Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. *Spine* 2002; 27 (23): 2674-83.

Deyo R A, Nachemson A, Mirza S K. Spinal-fusion surgery—the case for restraint. *N Engl J Med* 2004; 350 (7): 722-6.

Ekselius L. Effect of personality disorders on treatment outcome of surgery for low back pain. *J Musculoskeletal Pain* 1996; 4: 87-96.

Eriksen J, Jensen M K, Sjøgren P, Ekholm O, Rasmussen N K. Epidemiology of chronic non-malignant pain in Denmark. *Pain* 2003; 106 (3): 221-8.

Fairbank J C, Couper J, Davies J B, O'Brian J P. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980; 66 (8): 271-3.

Fanuele J C, Birkmeyer N J O, Abdu W A, Tosteson T D, Weinstein J N. The impact of spinal problems on the health status of patients: have we underestimated the effect? *Spine* 2000; 25 (12): 1509-14.

Hägg O, Fritzell P, Nordwall A, The Swedish Lumbar Spine Study Group. Characteristics of patients with low back pain selected for surgery: a comparison with the general population reported from the Swedish lumbar spine study. *Spine* 2002; 27 (11): 1230-1.

Hägg O, Fritzell P, Ekselius L, Nordwall A, The Swedish Lumbar Spine Study Group. Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish lumbar spine study. *Eur Spine J* 2003; 12: 22-33.

Hozack W J, Rothman R H, Albert T J, Balderston R A, Eng K. Relationship of total hip arthroplasty outcomes to other orthopaedic procedures. *Clin Orthop* 1997; (344): 88-93.

Keller S J, Soule D N, Singer D E, Deyo R A. Relationship between rates and outcomes of operative treatment for lumbar disc herniation and spinal stenosis. *J Bone Joint Surg (Am)* 1999; 81 (6): 752-62.

Leidy N, Revicki D, Genesté B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value in Health* 1999; 2: 113-27.

Lipson S J. Spinal-fusion surgery – advances and concerns. *N Engl J Med* 2004; 350 (7): 643-4.

Lucht U. The Danish Hip Arthroplasty Register. *Acta Orthop Scand* 2000; 71 (2): 111-21.

Mendelson G. Compensation, pain complaints, and psychological disturbance. *Pain* 1984; 20: 169-77.

Nachemson A. Back pain: delimiting the problem the next millennium. *Int J Law Psychiatry* 1999; 22 (5-6): 473-90.

Nachemson A, Zdeblic T A, O'Brian J P. Lumbar disc disease with dicogenic pain. What surgical treatment is most effective? *Spine* 1996; 21 (15): 1835-8.

- Rush A J, Polatin P, Gatchel R J. Depression and chronic low back pain: Establishing priorities in treatment. *Spine* 2000; 25: 2566-71.
- Ware J E Jr, Sherbourne C D. The MOS 36-item short-form health survey (SF-36) I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-83.
- Witting H, Turk D C, Carr DB, Sukiennik A, Rogers W. Comparison of the redundancy, reliability, and responsiveness to changes among SF-36, Oswestry Disability Index, and Multidimensional Pain Inventory. *Clin J Pain* 2004; 20 (3): 133-42.