# Outcome after total hip arthroplasty

## Part I. General health evaluation in relation to definition of failure in the Swedish National Total Hip Arthroplasty register

Peter Söderman, Henrik Malchau and Peter Herberts

Department of Orthopedics, Sahlgrenska University Hospital/Sahlgrenska, SE-413 45 Göteborg, Sweden.
Tel +46 31–342 18 39. Fax –82 55 99
Submitted 99-10-17. Accepted 00-02-03

**ABSTRACT – The Swedish National Total Hip Arthroplasty register, which was started in 1979, is one of the oldest national quality registers in the world and consists of over 160,000 primary operations and 12,500 revisions. The register identifies risk factors for poor outcomes related to the patient, implant and surgical techniques. The end-point for failure, i.e., revision, does not provide information about the patient's general health after the primary procedure. The aim of this study was twofold. First, to validate the end-point for failure in the Swedish National Total Hip Arthroplasty register and secondly, to study general health after total hip arthroplasty. We validated the outcome of 1,056 primary THRs randomly selected from the Discharge register in Sweden by comparing the data to the Swedish THA register. These patients had answered the SF-36 and Nottingham Health Profile questionnaires. By comparing the clinical outcome, measured as general health, with the results obtained from the register, we evaluated the importance of the end-point for failure. We found that the end-point was useful, but further evaluations are desirable. Patients operated on with hip replacement do very well up to 10 years postoperatively and those who are not revised have good general health. The findings in this study can be used as a reference for others as it shows results from a national register, with a random selection of the study cohort.**

■

The Swedish National Total Hip Arthroplasty register (the Swedish THA register), initiated 20 years ago, describes the epidemiology of total hip replacement in Sweden. It identifies risk factors for poor outcomes related to the patient, implant and surgical techniques (Herberts and Malchau 1997). All hospitals in Sweden provide the register with information about the amount and type of primary total hip arthroplasty and revision of hip prostheses performed annually (Malchau and Herberts 1994). Information in the register is disseminated yearly back to the clinics with reports of their own results of primary and revision surgery. These reports contain survival analysis of different types of implants and techniques. The end-point for failure is revision, that is, exchange or extraction of at least one part of the prosthesis. This is a precise but blunt instrument but does not provide information about the unrevised patient's general health (Malchau et al 1993). In the present study, we validated the end-point for failure in the Swedish THA register and measured the "precision" of this end-point by assessing the general health of unrevised patients after total hip arthroplasty done in the routine daily work.

## Patients and methods

We randomly selected 1,056 patients (number of patients on power analysis) operated on with primary hip prosthesis in Sweden, between 1986 and 1995, from the Discharge register (the Swedish National Board of Health and Welfare) (Söderman and Malchau 1997). The patients selected were checked in the Swedish Register of Deaths to ensure inclusion of only living patients. The patients received a short questionnaire concerning any re-

operations, pain and overall satisfaction. They were also given two self-administered general health questionnaires, the Nottingham Health Profile (NHP) and the SF-36. Two more letters were mailed to those who did not answer within 3 weeks. Patients who still did not reply were phoned and asked to send in the questionnaire.

### *Statistics*

We used the computerized statistical program SPSS (Chicago). Total score, domain scores, mean, standard deviation (95% CI), minimum value and maximum value (range) were calculated for all patients. The material was also divided into patients with one affected hip (Charnley category A), both hips affected (B), multiple joints affected or other disease that impairs gait (C) and gender was noted. The Mann-Whitney U-test was used to analyze the significance of the data.

For the 10-year survival analysis, we used logistic regression and the results were compared to the Swedish THA register. Values for clinical failure were empirically estimated as 50, 60 or 70 points. Clinical failure in this study was thus defined as a revised patient or a patient scoring less than 50, 60 or 70 points, giving four alternatives.

To see whether NHP and SF-36 measured what they were supposed to, we calculated the Spearman correlation between the same domains (convergent validity) and different domains (divergent validity). The hypothesis was that the same domains should have higher correlations with one another than with different domains. The internal consistency reliability of NHP and SF-36 was determined with the Cronbach (1951) alpha index.

### *Score systems*

The Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) has been used to study health-related quality of life (HRQL) (Brazier et al. 1992, Ware and Sherbourne 1992). The instrument consists of 36 questions divided into 8 domains: physical function, social function, role-emotional, role-physical, bodily pain, general health, mental health and vitality. The SF-36 is the most frequently used health-status measure in Northern America and has been used after total joint replacement in the hip and knee (Ware and Sherbourne 1992, Ritter et al. 1995, Martin et al. 1997). Sullivan et al. (1995) translated the test into Swedish and tested its validity and reliability. The raw score was transformed into a 0–100 scale (transformed scale score), as recommended by the Swedish manual. A high value indicated a better result.

The Nottingham Health Profile (NHP) is a self-administered general instrument used to study quality of life after treatments such as total hip replacement (Wiklund and Romanus 1988). The NHP has high validity and reliability (Hunt et al. 1980, 1981a, b). Wiklund and Romanus (1988) developed the Swedish version. The test consists of two parts with a total of 45 yes or no answers. There are 38 questions in Part 1 concerning patient difficulties, and it is divided into 6 domains: emotional reaction, sleep, energy, pain, physical mobility and social isolation. The items are weighted and each dimension yields a value between 0 and 100. Part 2 consists of 7 statements that reflect the frequency of problems in various areas of life: occupation, housework, social life, family life, sexual function, hobbies and holidays (McKenna et al. 1981). To avoid negative values in the validity test, the score was inverted, so that maximum health yielded 100 points.

## Results

The response rate for NHP and SF-36 was 93%. There were no major differences due to gender even though men received higher scores than women. 496 (46%) were men and the mean age was 76 (SD 9.3) years for men and 75 (SD 11) years for women. 86% were operated on for arthrosis, 3% for arthritis and 2% for sequelae after hip fracture. The total score for NHP increased from 15 to 24 up to 10 years postoperatively and for SF-36, the total score decreased from 70 to 60 during the same period (Table 1). The standard deviations for domain and total score (95% confidence interval), were 7.3–39 for NHP and 21–44 for SF-36. The results differed in patients with 1 affected hip than those with multiple joint disease (Table 2). This was especially true of NHP. Compared to a normal population, patients operated on with hip replacements do well, except for the functional domains (Figure 1).

**Table 1. Mean domains and total scores for NHP and SF-36 2–10 years postoperatively. High values in the SF-36 and low in the NHP indicate optimal health**

| Year postop. | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| *NHP* | | | | | | | | | |
| n | 115 | 110 | 94 | 144 | 127 | 92 | 92 | 103 | 104 |
| pain | 17 | 20 | 21 | 21 | 20 | 23 | 24 | 24 | 25 |
| energy | 20 | 22 | 26 | 32 | 18 | 25 | 32 | 31 | 35 |
| sleep | 16 | 23 | 20 | 19 | 17 | 23 | 21 | 25 | 26 |
| physical motion | 22 | 22 | 22 | 24 | 21 | 27 | 27 | 28 | 35 |
| emotional reaction | 7.7 | 11 | 12 | 10 | 8.8 | 9.2 | 16 | 14 | 15 |
| social isolation | 4.9 | 5.0 | 7.9 | 7.3 | 3.8 | 5.4 | 10 | 7.3 | 9.1 |
| total | 15 | 17 | 18 | 19 | 16 | 18 | 22 | 22 | 24 |
| *SF-36* | | | | | | | | | |
| n | 111 | 109 | 92 | 143 | 125 | 90 | 89 | 98 | 101 |
| physical function | 58 | 52 | 52 | 54 | 53 | 51 | 45 | 47 | 42 |
| role-physical | 54 | 50 | 50 | 49 | 53 | 47 | 43 | 41 | 41 |
| bodily pain | 73 | 63 | 65 | 67 | 67 | 65 | 64 | 62 | 58 |
| general health | 66 | 66 | 66 | 65. | 67 | 65 | 64 | 62 | 62 |
| vitality | 66 | 62 | 63 | 60 | 64 | 61 | 56 | 58 | 54 |
| social function | 86 | 82 | 81 | 82 | 84 | 83 | 79 | 80 | 77 |
| role-emotional | 67 | 63 | 66 | 65 | 69 | 64 | 59 | 62 | 58 |
| mental health | 81 | 79 | 76 | 78 | 81 | 78 | 75 | 77 | 75 |
| total | 70 | 66 | 65 | 66 | 67 | 65 | 63 | 63 | 60 |

**Table 2. Mean domains and total scores for Charnley categories A, B and C. High values in the SF-36 and low in the NHP indicate optimal health**

| | 2 years | | | 5 years | | | 10 years | | |
|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | A | B | C | A | B | C |
| *NHP* | | | | | | | | | |
| n | 52 | 18 | 43 | 38 | 28 | 75 | 21 | 13 | 53 |
| pain | 6.2 | 18 | 27 | 7.8 | 11 | 32 | 5.9 | 17 | 33 |
| energy | 8.5 | 212 | 31 | 14 | 22 | 45 | 24 | 31 | 40 |
| sleep | 11 | 13 | 20 | 13 | 14 | 25 | 15 | 31 | 27 |
| physical motion | 9.3 | 24 | 35 | 9.6 | 12 | 36 | 21 | 19 | 46 |
| emotional reaction | 2.9 | 6.2 | 14 | 3.4 | 7.6 | 15 | 13 | 12 | 19 |
| social isolation | 2.2 | 6.3 | 7.3 | 3.5 | 5.7 | 9.9 | 7.4 | 5.4 | 12 |
| total | 6.7 | 15 | 22 | 8.6 | 12 | 27 | 14 | 19 | 29 |
| *SF-36* | | | | | | | | | |
| n | 52 | 17 | 42 | 38 | 28 | 74 | 20 | 13 | 51 |
| physical function | 74 | 55 | 42 | 69 | 64 | 42 | 47 | 50 | 36 |
| role-physical | 66 | 53 | 42 | 69 | 63 | 34 | 66 | 54 | 28 |
| bodily pain | 88 | 66 | 58 | 84 | 72 | 55 | 86 | 68 | 47 |
| general health | 71 | 68 | 58 | 74 | 68 | 60 | 72 | 65 | 58 |
| vitality | 74 | 65 | 60 | 75 | 70 | 49 | 68 | 58 | 45 |
| social function | 92 | 74 | 84 | 90 | 84 | 767 | 84 | 83 | 70 |
| role-emotional | 73 | 73 | 57 | 75 | 72 | 57 | 73 | 72 | 43 |
| mental health | 87 | 78 | 75 | 89 | 83 | 71 | 80 | 79 | 70 |
| total | 78 | 68 | 60 | 78 | 71 | 57 | 77 | 65 | 51 |

10 years postoperatively, the Swedish National Total Hip Arthroplasty register showed a 93% survival rate (Figure 2). The clinical survival, defined as a revised patient or patient who had scored less than 50, 60 or 70 points in the total score for SF-36 was 35–57% after 10 years (Figure 2). The corresponding survival rate for NHP was 58–86% (Figure 3). The survival rate based on the logistic

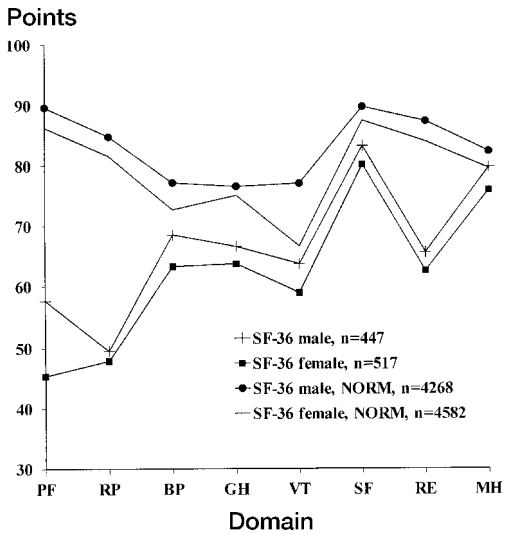


Figure 1. Mean domain score for SF-36 2–10 years postoperatively, compared to a normal (NORM) population (Swedish manual for SF-36). Physical function (PF), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social function (SF), role-emotional (RE) and mental health (MH)

regression analysis shows the probability for clinical failure or if the patients have been revised. It does not tell how good or bad the patients' general health is. As shown above, the patients in this study had scored close to those of a normal population.

There was no difference in general health between regional, county and rural hospitals as regards survival rates, based on the analysis of NHP and SF-36.

The correlation between the same domains in NHP and SF-36 was significant at the 1% level. The same domains, such as pain, showed higher correlations with each other (Spearman's rho 0.76) than with different domains, such as function (Spearman's rho 0.64). This finding indicates high convergent and divergent validity for NHP and SF-36. The questionnaires also showed high reliability (Cronbach's alpha 0.43–0.94), except for the general health and social function domains in SF-36, where the Cronbach's alpha indices were acceptable (0.43 and 0.50, respectively).

## Discussion

This study, like others, showed various clinical outcomes in patients with 1 hip affected and those with comorbidity that affects their gait (Garellick et al. 1998, Söderman and Malchau 1999). This is especially true when one uses general questionnaires for health evaluations. NHP and SF-36 can be used to evaluate outcome after hip replacement. General questionnaires should also be used

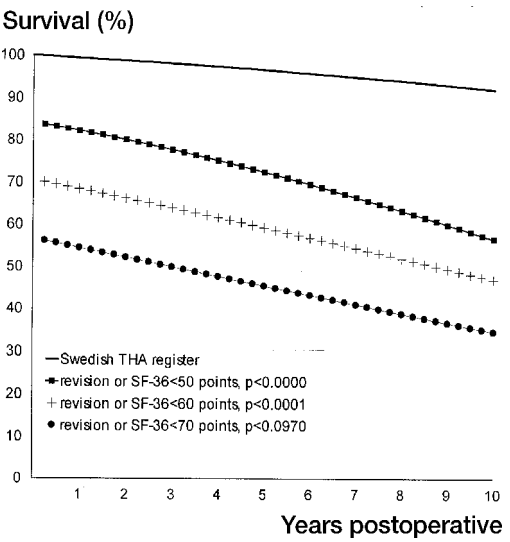


Figure 2. 10-year survival rate of all total hip arthroplasties in the Swedish National Total Hip Arthroplasty register (n 93,852), compared to the SF-36 (n 911).

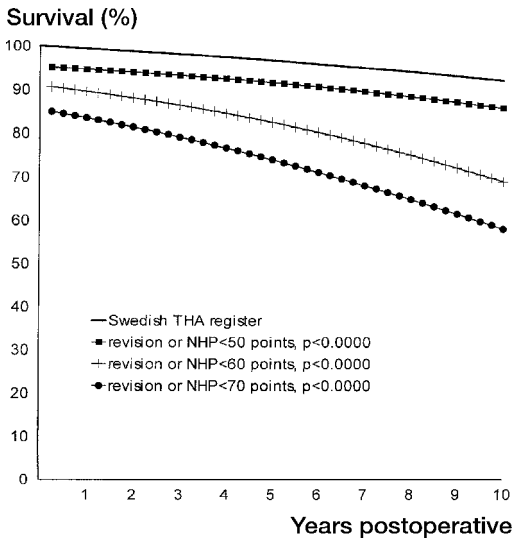


Figure 3. 10-year survival rate of all total hip arthroplasties in the Swedish National Total Hip Arthroplasty register (n 93,852), compared to the NHP (n 1,056).

to compare hip surgery with other medical or surgical interventions (Knahr et al. 1998). However, if one wants to study a new prosthesis design, one should use a disease-specific questionnaire (Pacault-Legendre and Courpied 1999, Söderman and Malchau 1999).

The end-point for failure in the Swedish THA register is reliable, but very strict. The outcome based on generic instruments, on the other hand, varies due to demographic bias and the definition of clinical failure. One obvious problem in comparing the results of NHP and SF-36 with the revision data in the Swedish THA register in this study was to determine the score that should be regarded as clinical failure. These levels have been arbitrarily estimated on the basis of each reference in the manuals for SF-36 and NHP, but for the older population (aged over 75 years), the scores for clinical failure in the SF-36 are probably less than 50 points. We tried to find a more exact level by examining a cohort from this study both clinically and radiographically (Söderman et al. 1999).

In summary, the end-point for failure in the Swedish THA register is a valuable definition of failure. Clinical failure depends on the level of failure, the type of score system used and comorbidity (demographic bias). In part II of the validation of the Swedish National Total Hip Arthroplasty register, it is reported that the survival rate will be compared to disease-specific tests, both self-administered (WOMAC) and staff-administered (Harris Hip Score). Patients operated on with hip replacement do very well up to 10 years postoperatively. Our material in this study can be used as a reference for other studies, since it shows results from a national register with a random selection of the study cohort.

We thank the Volvo Research Foundation, the Swedish National Board of Health and Welfare, the Gothenburg Medical Association, the Dr Felix Neuberg Foundation, the Greta & Einar Asker Foundation and the Hjalmar Svensson Foundation for financial support and A. Oden for statistical advice.

Brazier J E, Harper R, Jones N M B, O'Cathain A, Thomas K J, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305: 160-4.

Cronbach L J. Coefficient alpha and the internal structure of tests. Psychometrika 1951; 3: 297-334.

Garellick G, Malchau H, Herberts P. Specific or general health-outcome measures in the evaluation of total hip replacement. A comparison between the Harris Hip Score and the Nottingham Health Profile. J Bone Joint Surg (Br) 1998; 80 (4): 600-6.

Herberts P, Malchau H. How outcome studies have changed total hip arthroplasty practices in Sweden. Clin Orthop 1997; 344: 44-66.

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.

Hunt S M, McKenna S P, McEwen J, Williams J, Papp E. The Nottingham health profile: subjective health status and medical consultations. Soc Sci Med 1981a; 15A: 221-9.

Hunt S M, McKenna S P, Williams J. Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthrosis. J Epidemiol Community Health 1981b; 35: 297-300.

Knahr K, Kryspin Exner I, Jagsch R, Freilinger W, Kasparek M. Evaluating the quality of life before and after implantation of a total hip endoprosthesis. Z Orthop Ihre Grenzgep 1998; 136 (4): 321-9.

Malchau H, Herberts P. Medical data base. A Swedish program for outcome follow-up. Theme: National Total Hip Register 1979-1990. No. 1, National Board of Health and Welfare (ISSN 1104-862X), 1994.

Malchau M, Herberts P, Ahnfelt L. Prognosis of total hip replacement in Sweden. Follow-up of 92,675 operations performed 1978-1990. Acta Orthop Scand 1993; 64: 497-506.

Martin D P, Engelberg R, Agel J, Swiontkowski F. Comparison of the musculoskeletal function assessment questionnaire with the Short form-36, the Western Ontario and McMaster Universities Osteoarthritis index, and the Sickness Impact Profile Health-Status Measures. J Bone Joint Surg (Am) 1997; 79 (9): 1323-35.

McKenna S P, Hunt S M, McEwen J. Weighting the seriousness of perceived health problems using Thurstone's method of paired comparisons. J Epidemiol 1981; 10 (1): 93-7.

Pacault-Legendre P, Courpied J P. Survey of patient satisfaction after total arthroplasty of the hip. Int Orthop (SICOT) 1999; 23: 23-30.

Ritter M A, Albohm M J, Keating M, Faris P M, Meding J B. Comparative outcomes of total joint arthroplasty. J Arthroplasty 1995; 10 (6): 737-41.

Sullivan M, Karlsson J, Ware J R. The Swedish SF-36 health survey - I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 41 (10): 1349-58.

Söderman P, Malchau H. Outcome measurements in total hip replacement surgery (THR). Outcome measuring. Spri Tryck 1997; 310: 89-95.

Söderman P, Malchau M, Herberts P. Outcome after total hip arthroplasty. Part II – disease specific questionnaires and the Swedish National Total Hip Arthroplasty Register. Submitted to Acta Orthop Scand, January 1999.

Ware J E, Sherbourne C D. The MOS 36-Item Short-Form Health Survey (SF-36). Med Care 1992; 30 (6): 473-83.

Wiklund I, Romanus B. Nottingham health profile. Livskvalitetsbedömning hjälp vid operationsprioritering. Läkartidningen 1988; 85 (38): 3060-1.