Osteoarthritis and Cartilage (2001) **9,** 137–146
© 2001 OsteoArthritis Research Society International
doi:10.1053/joca.2000.0369, available online at http://www.idealibrary.com on IDEAL®
1063–4584/010/020137+10 $35.00/0

Osteoarthritis and Cartilage
Journal of the OsteoArthritis Research Society International

A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery

C. J. M. Bachmeier*, L. M. March†, M. J. Cross*, H. M. Lapsley‡, K. L. Tribe*, B. G. Courtenay* and P. M. Brooks§ for the Arthritis Cost and Outcome Project Group**
**Departments of Medicine and Orthopaedics, St Vincent's Hospital, University of New South Wales, Australia*
†Departments of Rheumatology and Orthopaedics, Royal North Shore Hospital, University of Sydney, Australia
‡School of Health Services Management, University of New South Wales, Australia
§Office of the Executive Dean (Health Sciences), University of Queensland, Australia
||Australian Institute of Musculoskeletal Research, Sydney, Australia
¶North Sydney Orthopaedic & Sports Medicine Centre, Australia

Summary

Objective: The aims of this study were to assess changes in physical function and quality of life with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and the instrument of the Medical Outcomes Study SF-36 Health Survey (MOS SF-36), respectively, in patients undergoing hip and knee joint replacement surgery and to compare the responsiveness of these two outcome measures 1 year after surgery.

Design: One hundred and ninety-four patients with osteoarthritis (OA knee 108, OA hip 86) admitted to four hospitals in Sydney were followed over a period of 1 year at 3 monthly intervals.

Results: WOMAC measures improved significantly after 1 year for OA hip and OA knee: there was reduction in pain of 71% and 53%, reduction in stiffness of 55% and 43% and improvement in physical function of 68% and 43%, respectively. MOS SF-36 measures in those having hip surgery improved significantly for pain (222%), physical function (247%), physical role functioning (402%), general health (110%), vitality (143%), social functioning (169%) and mental health (114%). For those in the knee surgery group, significant improvement was seen for pain (175%), physical function (197%), physical role functioning (275%), vitality (125%) and social functioning (119%). The WOMAC was a more responsive measure than the MOS SF-36.

Conclusion: WOMAC and MOS SF-36 detect significant and clinically meaningful changes in outcome after hip and knee replacement. WOMAC requires a smaller sample size and is more responsive in the short term. For a follow-up longer than 6 months MOS SF-36 provides additional information. The improvement in outcomes following hip joint surgery were significantly greater than those following knee surgery.

Key words: Hip/knee surgery, Osteoarthritis, WOMAC, MOS SF-36.

Introduction

Joint replacement has had a major impact on the management of osteoarthritis (OA). Economically it requires an important and increasing part of the health budget in every country where it has been investigated. In the U.S.A. more than 400 000 primary hip and knee arthroplasties are performed each year and the annual costs of these procedures were estimated to exceed US$ 10 billion[1]. In Australia AUS$ 13.5 million Commonwealth benefits were paid in the financial year 1997/98 for 17 000 hip and knee replacements (customized tables, Health Insurance Commission), compared to AUS$ 8.5 million in 1991/92 for 14 000 hip and knee replacements[2]. Expenditures on primary joint arthroplasty surgery will increase in the future with our aging population.

In clinical research, outcome instruments are of major importance. They have to be valid, reliable and responsive to change. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) is a multidimensional, self-administered outcome measure, which has been developed by Bellamy[3] for clinical trials in patients with hip or knee osteoarthritis. It probes for the dimensions of pain (5 items), stiffness (2 items) and physical function (17 items). It has been intensively validated[4] and shown to be a valid, reliable and responsive instrument. The Medical Outcomes

Received 13 May 1999; accepted 10 June 2000.
**Including: C. A. Cass*; M. R. J. Coolican†; M. J. Cross||¶; M. J. Neil*; L. A. Pinczewski||¶; J. S. Quain*; F. W. Robertson*; S. J. Ruff†.
Sources of support: Australian National Health and Medical Research Council, Ciba-Geigy-Jubiläumsstiftung Basel Switzerland, Karger Fonds Basel Switzerland, Schweizerische Gesellschaft der Wissenschaften Basel Switzerland, Schweizerische Akademie der Wissenschaften in Basel, Switzerland and Schweizerische Gesellschaft für Rheumatologie in Switzerland.
Address correspondence to: Marita Cross, Level 9, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst NSW 2010, Australia. Tel: 0061 2 9295 8419; Fax: 0061 2 9295 8401.

Study SF-36 Health Survey (MOS SF-36) is a self-administered, generic health-related quality of life instrument. It was developed in the Rand Health Insurance experiment from the Rand Health Insurance long form. Quality of life is assessed by 36 items across eight dimensions (physical functioning, role/physical functioning, bodily pain, general health, vitality, social functioning, role/emotional functioning, mental health). It has been tested for its psychometric properties[5,6]. It has not been specially designed for patients with OA but seems to be applicable to a broad spectrum of diseases. Compared to the WOMAC it is a more general instrument with the advantage of the ability to assess pain, physical function and quality of life, which are the most important outcomes in clinical trials with osteoarthritic patients. The WOMAC and the MOS SF-36 have been recommended as valid outcome measures in OA research at the OMERACT III conference[7] and in the guidelines of the Osteoarthritis Research Society[8] to allow standardization and comparability between studies. In this study we wanted to quantify outcome after hip and knee replacement using these two recommended measures, the WOMAC and the MOS SF-36, and compare them prospectively. We hypothesized that the WOMAC would be the more efficient instrument, being disease-specific.

Method

STUDY POPULATION

This study is part of an ongoing large prospective trial assessing costs of arthritis. Nine orthopedic surgeons in four hospitals in Sydney, Australia (St Vincent's Public and Private Hospitals, the Centre for Bone and Joint Disease and Mater Misericordiae) provided operation waiting lists for total primary hip or knee joint replacement. Both privately insured and public (non-insured) patients were included. All patients with a diagnosis of OA and rheumatoid arthritis were eligible to enter the study. Waiting lists were checked fortnightly to monitor the continuing recruitment. Patients were asked by telephone to take part in the trial. In this paper only patients with OA were analysed. Most patients were visited at home by the research assistant for an interview after having signed a consent form, and were also seen on admission for surgery. Follow-up was by postal questionnaires or home visit if necessary. Follow-up home visits were made to four hip replacement and eight knee replacement patients. For two of the hip patients and two of the knee patients reading English was a problem, so the questions were read out to these people by the interviewer. For the remainder, the questionnaires were self-administered and the research assistant provided assistance where necessary.

ASSESSMENTS

Over a 12-month period the patients were mailed WOMAC and MOS SF-36 questionnaires at 3-monthly intervals. This analysis includes the first 12-month period of follow-up. Shortly before the questionnaires were due back patients were reminded by telephone calls. In case of failure to respond the patients were again reminded by telephone. Returned questionnaires were checked for missing answers and clarified by telephone if necessary. To minimize the drop-out rate patients were informed about the study progress with quarterly newsletters and personal contacts such as birthday cards.

ANALYSIS

Data were entered using Paradox software and then analysed using Statistical Package for Social Sciences (SPSS) for Windows version 6.0 program[9]. Both questionnaires were scored according to guidelines of their authors[10,11]. *P*-values ≤ 0.05 were considered significant. Parametric Students *t*-tests (independent and paired tests) were used to assess the significance of the mean differences before and after the operation, when scores were approximately normally distributed. The non-parametric test (Mann–Whitney U-Wilcoxon sum W test) was used for the samples that were not normally distributed. Normality of distribution was examined by standard diagnostic tests (Lillefors test, box plot, normal probability plot and histogram). The distribution of age, gender and duration of disease in the groups of hip and knee joint replacement were assessed by chi-squared tests (for the dichotomous variable, gender) and independent *t*-tests (for continuous variables). The change in the different domains was calculated as the difference of the pre-operative score to the follow-up score. The improvement in the hip replacement group compared to the knee replacement group was the comparison of the mean score difference of the pre-operative score and the follow-up score between the hip and the knee replacement groups. These two comparative analyses used paired tests limiting the calculations to those people who completed both questionnaires at each time point. These calculations required multiple comparisons. Adaptation of the significance level was not performed. The comparison of the improvement in the hip and knee replacement groups was further assessed using multivariate analysis to allow for age, gender and OA duration as potential confounders. The models were created by backward stepwise selection.

The estimates of the comparative responsiveness of the two instruments were calculated with the relative efficiency (RE) using the following formula)[12]:

$$\text{RE (WOMAC vs MOS SF-36)} = \frac{\text{t WOMAC}^2}{\text{t MOS SF-36}}$$

This comparison is related to pain and physical function, the only two dimensions measured in both questionnaires.

The MOS SF-36 scores on a 0–100 worst to best scale. The WOMAC scores on a 1–5 best to worst scale, which was transformed accordingly to the MOS SF-36 0–100 worst to best scale.

Results

From 24 February 1994 to 1 July 1996, 376 patients with OA on the waiting lists for hip or knee joint replacement were eligible to take part in the study. Of these, 124 patients did not take part in the study for the following reasons: 49 patients could not be contacted in time, 27 were not interested, 18 felt too old or unwell, the English of 12 patients was too poor, seven patients could not be contacted at the address provided by the surgeon, five patients did not have a telephone to be contacted by, four patients found the study too confusing and the operation was cancelled in two patients. There was no difference in age, sex or type of operation between those who provided pre- and post-operative questionnaires and those who did not take part in the study.

Table I
*Baseline characteristics of patients**

Characteristic	Hip replacement group (*N*=86)	Knee replacement group (*N*=108)	Entire surgical group (*N*=194)
Age (years)†	65 (11.5)	72 (7.0)	69 (9.9)
Duration of disease (years)‡	8.5 (9.4)	12.0 (11.9)	10 (11.1)
Female proportion (%)	39	61	52

*Providing at least one pre-operative and one post-operative WOMAC questionnaire.
†Mean (standard deviation).
‡Median (standard deviation).

Two hundred and fifty-two patients started the study, corresponding to a response rate of 67.3%. Thirty-six patients provided pre-operative questionnaires and withdrew before providing post-operative questionnaires for the following reasons: two patients died, in seven patients the operation was cancelled, two patients moved from their address and left no forwarding details and 25 patients denied further participation in the study. These 36 patients (14.3%) did not differ regarding sex and type of joint replacement operation but were older compared to the 194 patients in the analysis ($t=-2.506$, $P=0.013$, mean age of patients lost to follow-up and participants: 70±9.9 years and 74±10.4 years, respectively). Twenty-two patients started to provide pre-operative data but their first post-operative questionnaire was answered after 1 July 1996.

For 194 patients (86 with OA of the hip and 108 with OA of the knee) at least one pre-operative and one post-operative WOMAC questionnaire was available. One hundred and seventy-three patients (73 with OA of the hip and 100 with OA of the knee) provided at least one pre-operative and one post-operative MOS SF-36 questionnaire in addition to the WOMAC questionnaires. Ninety-one per cent of hip questionnaires were completed within 4 weeks prior to surgery, another 1.2% up to 5 weeks and 7.8% were longer than 5 weeks. For the knee patients, 88% were completed within 4 weeks pre-surgery, another 4.5% within 5 weeks and 7.5% in excess of 5 weeks. Where questionnaires were completed in excess of 5 weeks prior to surgery, these people were sent questionnaires, but did not complete them immediately pre-operatively. The mean time to surgery from pre-operative evaluation was 14.9 days for knee replacement patients and 14.2 days for hip replacement patients.

The overall cohort was followed for a total of 2031 person-months, with an average of 11 months. For the hip group, the total follow-up was 891 person-months (average 11 months) and for the knee group 1140 person months in total (average 10 months).

STUDY POPULATION CHARACTERISTICS

The study population characteristics of age, duration of disease and gender distribution are listed in Table I. In general, the patients in the hip replacement group were younger (t-value$=-5.27$, 130 df, $P<0.001$) and had a shorter duration of their disease (z-value$=-2.68$, $P=0.007$) than those undergoing knee replacement surgery. Pre-operatively all the domains in the WOMAC and the MOS SF-36 were comparable between the two surgical cohorts (WOMAC: pain $P=0.19$, stiffness $P=0.99$, physical function $P=0.52$; SF-36: physical function $P=0.78$, role physical function $z=-1.69$, pain $P=0.46$, general health $P=0.16$, vitality $P=0.54$, social function $z=-1.86$, role emotional function $P=0.68$, mental health $P=0.98$).

Four in the hip group had their other hip replaced and three in the knee group had their other knee replaced in the 12-month follow-up period. Those who had another joint replaced did not differ significantly when WOMAC and SF-36 scores were compared with those who had no further joint replacement at the 12-month follow-up. There was a low number of hospitalizations for adverse events most probably related to the primary arthroplasty (3% hip and 13% knee). These included 10 knee patients requiring manipulation under anesthesia, three with an infection in the operated joint and one knee patient and three hip patients with thromboembolic complications. The 12-month WOMAC and SF-36 measures showed a significant difference in WOMAC physical function only with those requiring readmission to hospital within 12 months of knee arthroplasty reporting worse function at 12-month follow-up. No significant differences were seen in the hip replacement group in WOMAC or SF-36 scores at 12 months between those requiring readmission and those who did not.

In addition, four knee replacement patients reported they required hospitalization for arthritis-related conditions not related to the primary replacement within the 12-month follow-up period (such as neck pain due to arthritis, removal of cartilage from other knee). These patients reported significantly worse WOMAC pain, stiffness and function at 12-month follow-up than all other knee replacement patients (which includes those who were admitted for another arthroplasty or for conditions likely to be related to their primary arthroplasty). No difference was seen in any scale of the SF-36.

IMPROVEMENT OVER TIME IN THE HIP REPLACEMENT GROUP ACCORDING TO WOMAC

The improvement over time in the patient group undergoing hip replacement surgery is shown in Fig. 1. In the hip group there was a reduction in the mean scores of pain, stiffness and physical function as assessed by the WOMAC of about 50% in the first 3 follow-up months after the operation. This improvement continued up to 1 year, when the improvement in physical function and pain from baseline was more than 60%. The changes were statistically significant at 3, 6, 9 and 12 months for each of the WOMAC dimensions.

IMPROVEMENT OVER TIME IN THE KNEE REPLACEMENT GROUP ACCORDING TO WOMAC

The patients undergoing knee replacement surgery showed similar results (Fig. 2). The pre-operative scores

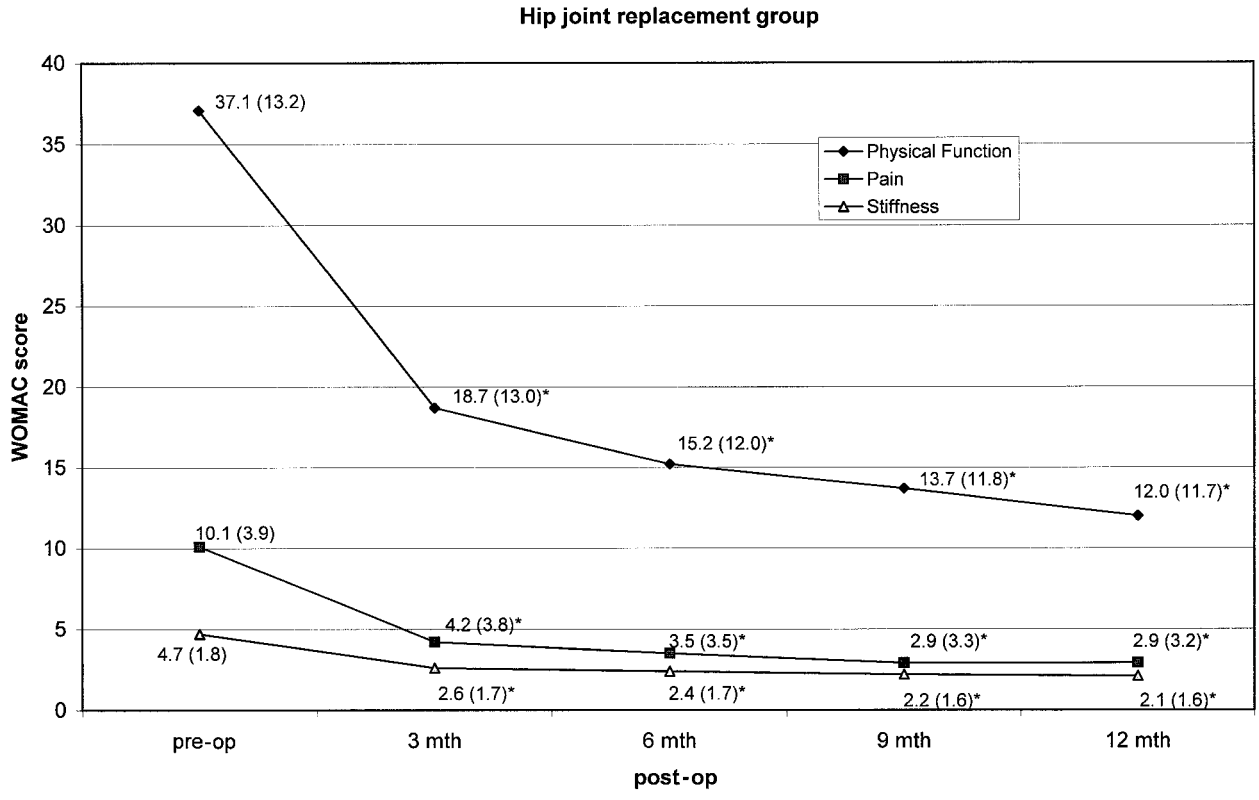

Fig. 1. WOMAC scores (means and standard deviations) in the patient group undergoing hip joint replacement surgery. *P<0.05. Lower scores indicate a better health state. Baseline N=86; 3 months N=82; 6 months N=64; 9 months N=61; 12 months N=43.

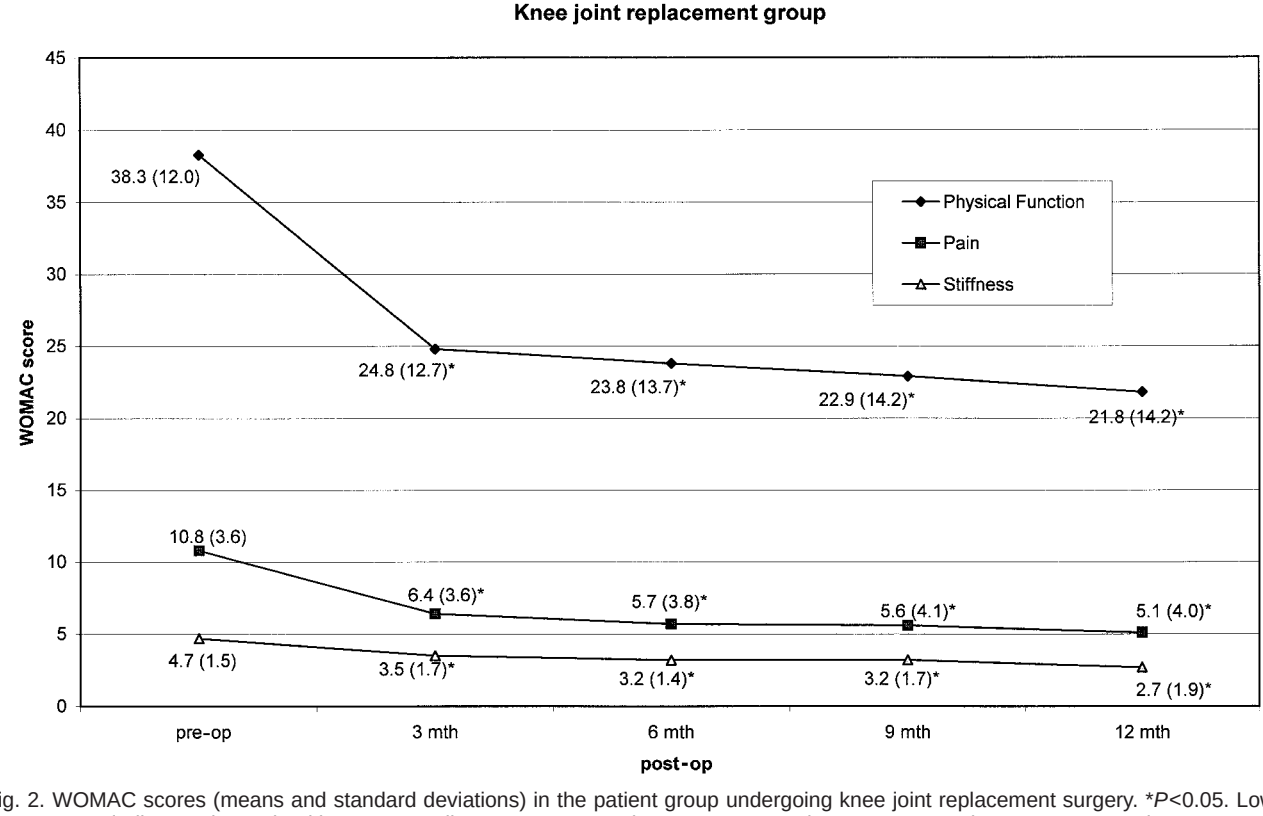

Fig. 2. WOMAC scores (means and standard deviations) in the patient group undergoing knee joint replacement surgery. *P<0.05. Lower scores indicate a better health state. Baseline N=108; 3 months N=101; 6 months N=84; 9 months N=67; 12 months N=48.

Table II
Improvement in the hip replacement group compared to the knee replacement group according to WOMAC

WOMAC domain	Mean differences (95% CI) between hip and knee change from baseline scores to post-operative follow-up			
	3 months	6 months	9 months	12 months
N	Hips: 79 Knees: 101	Hips: 62 Knees: 84	Hips: 67 Knees: 59	Hips: 42 Knees: 48
Pain	−1.3 (−2.7,0.04)	−0.9 (−2.4,0.6)	−2.0 (−3.6,−0.5)*	−2.6 (−4.6,−0.6)*
Stiffness	−0.9 (−1.5,−0.2)*	−0.8 (−1.4,−0.2)*	−1.1 (−1.7,−0.5)*	−0.8 (−1.6,0.1)
Physical function	−4.3 (−8.7,0.0)*	−5.2 (−9.4,−0.9)*	−7.9 (−12.7,−3.0)*	−10.8 (−17.1,−4.5)*

*$P<0.05$.
A negative score means a greater improvement in the hip replacement group.
CI=Confidence intervals.

for pain, stiffness and physical function as assessed by the WOMAC improved after operation. The improvement compared with baseline was statistically significant at 3, 6, 9 and 12 months. After 1 year the reduction in pain and stiffness and the improvement in physical function was each about 50%, respectively.

IMPROVEMENT IN THE HIP REPLACEMENT GROUP COMPARED TO THE KNEE REPLACEMENT GROUP ACCORDING TO WOMAC

The improvement in the hip joint replacement group compared to the knee joint replacement group is shown in Table II. The mean improvement in pain, stiffness and physical function between baseline and the follow-ups were more pronounced in patients who had undergone a hip replacement compared with patients who had knee joint replacement. Using univariate analysis the improvement was statistically significantly greater at all follow-ups regarding physical function, at follow-ups 9 and 12 months regarding pain and at follow-ups 3 and 6 months regarding stiffness. Multivariate analysis adjusting for age, sex and duration of disease confirmed a greater improvement in outcome for hip surgery patients at 9 months follow-up (pain $P=0.008$, stiffness $P=0.0001$ and physical function $P=0.001$, respectively) and at 12 months follow-up (pain $P=0.02$, stiffness $P=0.02$ and physical function $P=0.001$, respectively).

IMPROVEMENT OVER TIME IN THE HIP REPLACEMENT GROUP ACCORDING TO MOS SF-36

The improvement over time in patients undergoing hip joint replacement is presented in Fig. 3(a) and (b) (the domains of physical function, physical role function, bodily pain and general health and the domains of vitality, social function, emotional role function and mental health, respectively). In the hip replacement group all eight domains of the MOS SF-36 improved after the joint operation. The improvement in physical functioning, physical role functioning, pain, vitality, social functioning and mental health compared to baseline was statistically significant at 3, 6, 9 and 12 months. The improvement in general health was statistically significant at 9 and 12 months, whereas the improvement in emotional role functioning was statistically significant only at 9 months. The relative improvement after 1 year was most pronounced for physical functioning (247%) and physical role functioning (402%). These two latter dimensions showed statistically significant improvement also from 3–6 months. The quarterly improvement for pain, vitality, social functioning and mental health was statistically significant only for the first quarter postoperatively.

IMPROVEMENT OVER TIME IN THE KNEE REPLACEMENT GROUP ACCORDING TO MOS SF-36

Figure 4(a) (physical function, physical role function, bodily pain and general health) and (b) (vitality, social function, emotional role function and mental health) present the improvement over time in those undergoing knee joint replacement surgery. In the knee replacement group all domains of the questionnaire improved post-operatively except general health, which was not influenced by the surgical intervention. The improvement in physical functioning and pain was statistically significant at every follow-up, whereas the improvement in physical role functioning, vitality and social functioning was significant from 6 months on. As in the hip group the relative improvement at one year was most pronounced for the physical functioning (197%) and the physical role functioning (275%).

IMPROVEMENT IN THE HIP REPLACEMENT GROUP COMPARED TO THE KNEE REPLACEMENT GROUP ACCORDING TO MOS SF-36

Table III shows the change in the outcome of patients undergoing hip joint replacement surgery compared to those undergoing knee joint replacement surgery. Compared with patients after knee joint replacement, patients in the hip replacement group had a greater improvement for all domains at all follow-ups except regarding emotional role functioning at 12 months as assessed by the MOS SF-36. In univariate and multivariate analysis the greater improvement was significant at most of the quarterly follow-ups regarding physical function, pain and vitality. Significantly greater improvement in the hip group regarding physical role functioning using univariate analysis was confirmed by multivariate analysis only at three months post-operatively, whereas the greater improvement in mental health in the hip group was significant at 9 and 12 months after adjusting for age, sex and OA duration.

RELATIVE EFFICIENCY OF THE WOMAC COMPARED TO MOS SF-36

The WOMAC had a greater relative efficiency for pain and physical function at all follow-ups except for physical function in the knee replacement group at 12 months, where both questionnaires are similarly efficient (Table IV).

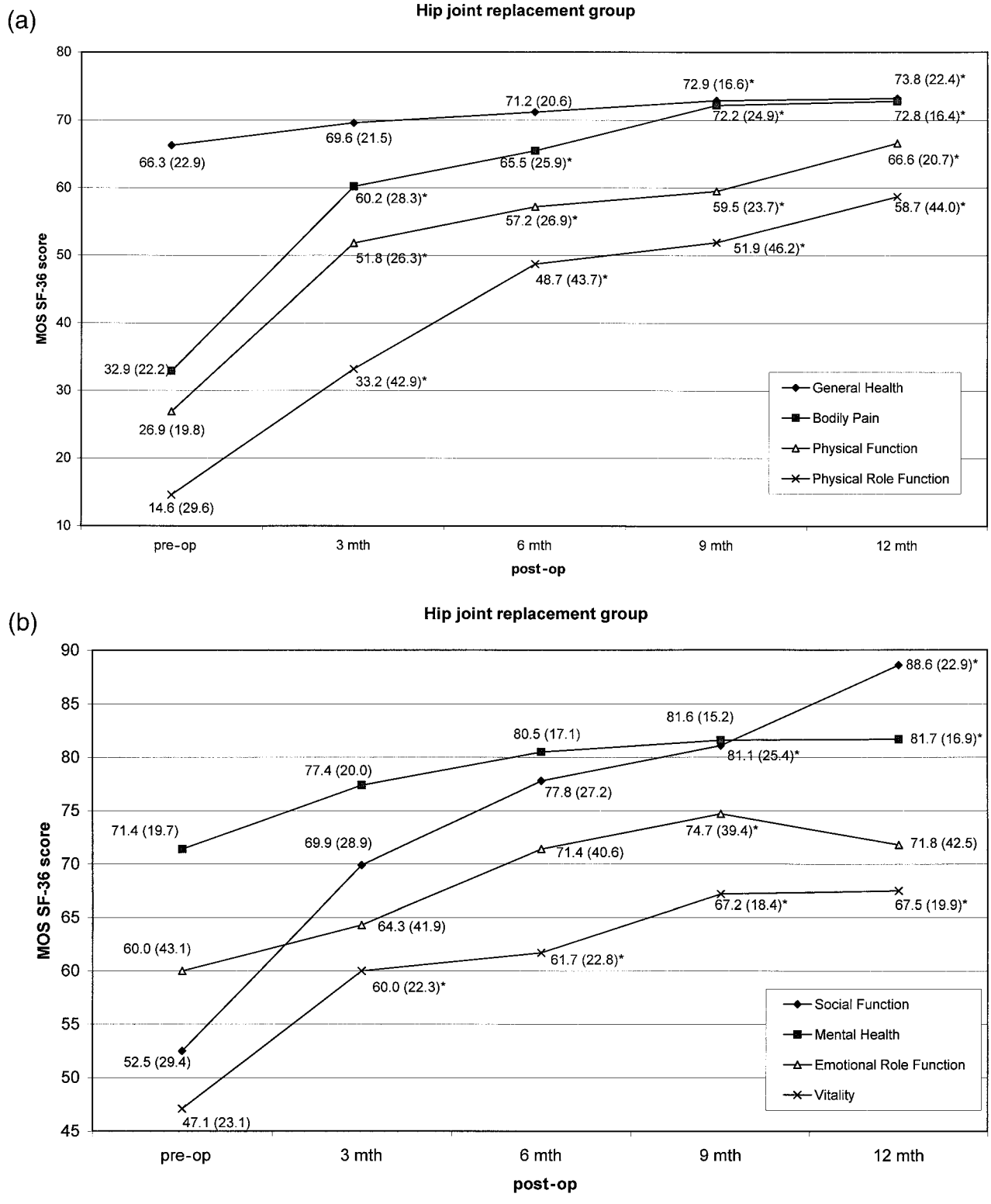

Fig. 3. (a) The MOS SF-36 scores of the domains physical function, physical role function, bodily pain and general health (means and standard deviations) in the patient group undergoing hip joint replacement surgery. *P<0.05. Higher scores indicate a better health state. Baseline N=73; 3 months N=71; 6 months N=57; 9 months N=54; 12 months N=53. (b) The MOS SF-36 scores of the domains vitality, social function, emotional role function and mental health (means and standard deviations) in the patient group undergoing hip joint replacement surgery. *P<0.05. Higher scores indicate a better health state. Baseline N=73; 3 months N=71; 6 months N=57; 9 months N=54; 12 months N=53.

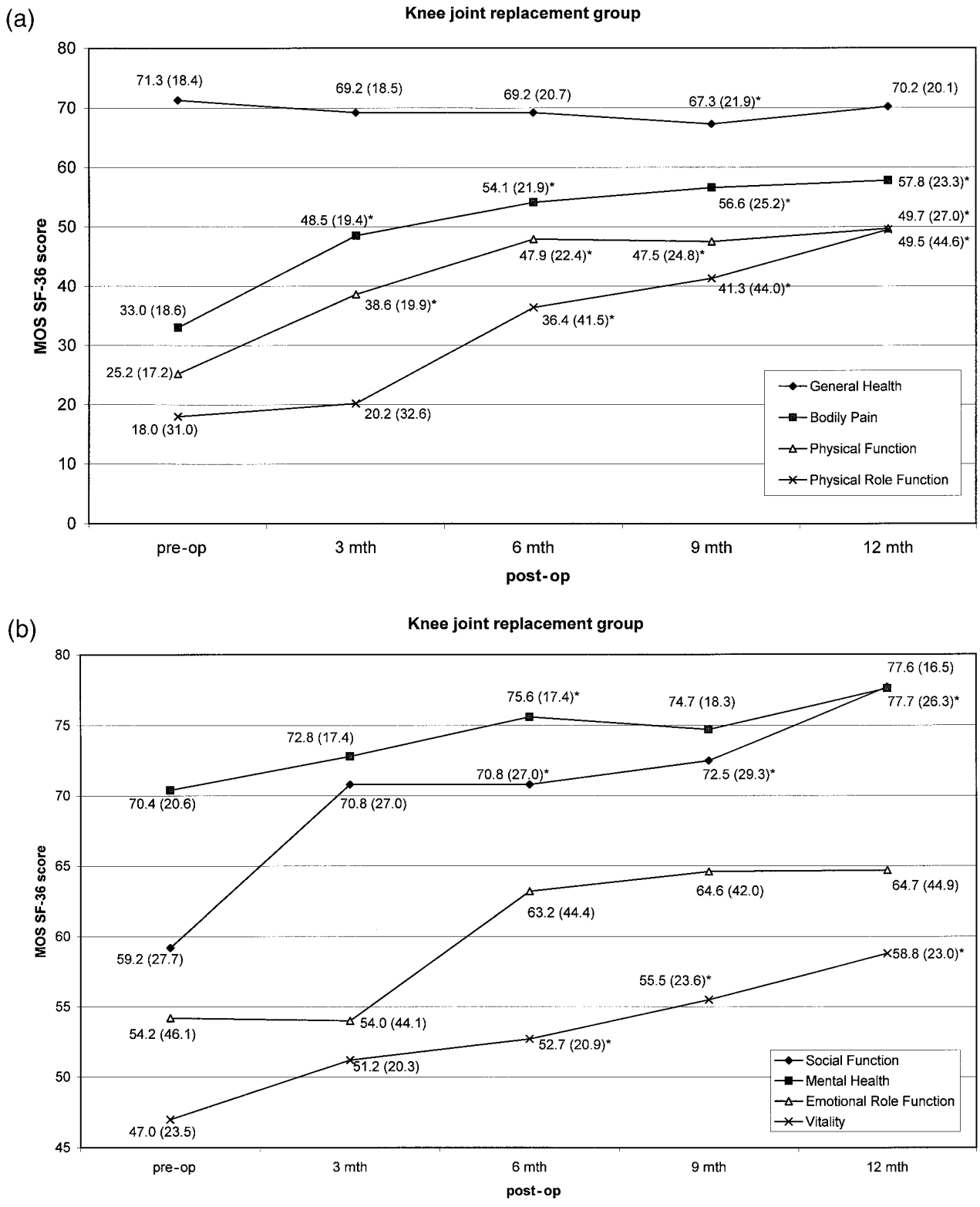

Fig. 4. (a) The MOS SF-36 scores of the domains physical function, physical role function, bodily pain and general health (means and standard deviations) in the patient group undergoing knee joint replacement surgery. *P<0.05. Higher scores indicate a better health state. Baseline N=100; 3 months N=89; 6 months N=77; 9 months N=66; 12 months N=51. (b) The MOS SF-36 scores of the domains vitality, social function, emotional role function and mental health (means and standard deviations) in the patient group undergoing knee joint replacement surgery. *P<0.05. Higher scores indicate a better health state. Baseline N=100; 3 months N=89; 6 months N=77; 9 months N=66; 12 months N=51.

Table III
Improvement in the hip replacement group compared to the knee replacement group according to SF-36

SF-36 domain	Mean difference (95% CI) between hip and knee change from baseline score to post-operative follow up			
	3 months	6 months	9 months	12 months
N	Hips: 66 Knees: 88	Hips: 52 Knees: 76	Hips: 49 Knees: 63	Hips: 42 Knees: 45
Physical function	10.7 (3.6,17.8)*	4.3 (−3.6,12.2)	8.4 (−0.2,17.0)	11.9 (1.6,22.3)*
Physical role function	17.3 (3.7,30.8)*	16.1 (0.5,31.7)*	17.0 (0.8,33.3)*	9.2 (−10.2,28.7)
Bodily pain	12.3 (4.5,20.2)*	8.2 (−1.2,17.6)	18.4 (8.4,28.4)*	18.5 (7.1,30.0)*
General health	6.2 (−0.5,12.9)	5.5 (−1.2,12.2)	10.9 (3.1,18.7)*	5.3 (−2.8,13.4)
Vitality	7.8 (−0.5,16.1)	8.0 (0.2,15.9)*	16.6 (8.3,24.9)*	11.2 (2.2,20.2)*
Social function	3.7 (−10.3,17.8)	3.9 (−8.5,16.3)	7.7 (−6.4,22.0)	9.5 (−5.2,24.3)
Emotional role function	6.2 (−12.2,24.6)	0.5 (−18.5,19.6)	5.2 (−16.2,26.5)	−1.2 (−22.7,20.3)
Mental health	3.8 (−2.3,9.9)	1.7 (−4.5,7.8)	8.2 (0.7,15.5)*	7.4 (−0.2,15.0)

*$P<0.05$.
A positive score means a greater improvement in the hip replacement group.
CI=Confidence intervals.

The greater relative efficiency of the WOMAC decreased over time for both surgical groups regarding pain and physical function. The relative efficiency of the WOMAC was most pronounced for pain assessment in the knee joint replacement group.

Discussion

Eighty-six patients undergoing hip joint replacement and 108 patients undergoing knee joint replacement were assessed pre-operatively and at 3-monthly follow-up intervals post-operatively with WOMAC and MOS SF-36 questionnaires.

As ascertained by the WOMAC both patient groups experienced significant improvement in pain, stiffness and physical function. After 1 year, pain was reduced to 29% and 47% of the pre-operative measures in the hip and knee group, respectively, stiffness was reduced to 45% and 57%, respectively and physical function increased by 68% and 43%, respectively. The greatest improvement was seen in both groups within the first 3 months post-operatively. After 3 months all measures continued to improve with the absolute values levelling off. A dramatic improvement by three months and smaller changes thereafter in these patients has been described previously[13].

As assessed by the MOS SF-36 both patient groups improved at all domains except that the general health of patients after knee replacement remained unaltered. Again, the greatest improvement in patients undergoing hip or knee replacement was seen within the first 3 months post-operatively. It seems a paradox that patients' perception of their general health was not influenced after knee replacement although significant changes in health status could be demonstrated. An unchanged general health status was seen previously in other 2-year follow-up studies looking at total hip and knee arthroplasties assessed by the MOS SF-36[14,15]. This may be explained by the fact that even after a very successful operation pain and restrictions in daily life remain post-operatively, which are perceived by patients strongly enough to rate their general health as insufficient. Perhaps this paradox reflects some limitations of a generic health questionnaire which does not include other important aspects of quality of life, such as sleep and sexual function. Correlations between pain reduction and functional independence and psychological well-being have been shown by others to be inconsistent[16].

Table IV
The relative efficiency for WOMAC vs MOS SF-36

Follow-up (months)	Pain		Physical function	
	Hip	Knee	Hip	Knee
0–3	1.43	1.94	1.73	1.73
0–6	1.54	1.92	1.40	1.17
0–9	1.42	1.84	1.48	1.24
0–12	1.13	1.74	0.99	1.30

Relative efficiency >1 if WOMAC is the more responsive measure.

Although patients in the hip and the knee replacement group experienced improvement in virtually all domains after their operation, the improvement was earlier and more pronounced in patients with hip replacement for all domains of the WOMAC and for physical function, pain, general health, vitality and mental health assessed by the MOS SF-36. The reason for these differences is not clear, as there were no statistically significant differences in the outcome measures at baseline and using multivariate analysis potential confounders such as age, sex and duration of disease were adjusted for. Within this 1 year of follow-up no patient had to undergo a revision of the study knee or hip arthroplasty or a previously implanted hip or knee replacement. It is possible that these differences between hip and knee replacement groups could be explained by a different rate of post-operative complications, which were slightly higher in the knee group. Comorbidities may also contribute; however, these were not assessed. A greater relative improvement in osteoarthritic hip vs knee replacement for pain and disability has been described elsewhere[17,18], including a 10-fold difference in Quality Adjusted Life Years (QALYs) gained[19] when hip replacement (4 QALYs gained) were compared to knee replacement (0.42 QALYs gained)[20].

As a generic health status measure, the SF-36 has been found to be more relevant and more responsive than the Sickness Impact Profile and is the preferable generic health status measure in patients undergoing hip replacement[21]. In this study the disease-specific WOMAC was a more responsive measure than the MOS SF-36. This was most pronounced in the assessment of pain in patients with knee

replacement. After 1 year the greater relative efficiency of the WOMAC regarding ascertainment of pain and physical function in both patient groups levelled somewhat compared to the first months of observations. This may be important for short-term studies with a duration of less than 1 year, where using the WOMAC questionnaire requires a smaller sample size. In agreement with the results of this study the WOMAC has been found to be the more sensitive measure in detecting knee disabilities compared to the MOS SF-36[22]. This is not surprising, as the WOMAC has been developed as a specific outcome measure for patients with OA of the lower limb.

Applying the MOS SF-36 questionnaire provided additional information on the patients' quality of life as expressed by physical role functioning, general health, vitality, social functioning, emotional role functioning and mental health. This information is clinically important and patient-relevant, as well as the more traditional outcomes of pain, stiffness and physical function. Quality of life is a major determinant of the patients' global assessment of satisfaction with their health status. Interestingly, quality of life was slower to improve after the operation and the relative changes were more subtle compared to the relative improvement in physical outcome measures. Similarly, a less dramatic improvement of psychological function compared to pain was noted in a cohort study of patients undergoing hip replacement surgery with a 6–8 week follow-up[23] which may reflect a slower adaptation of psychological parameters compared to physical issues in a new life situation.

The strengths of the study are its prospective design, high participation and follow-up rate reducing well known sources of biases of retrospective and cross-sectional studies. The characteristics of the study population from multiple orthopedic surgeons and its source both in private and public hospitals should allow the generalizability of the results. Comparable results have been reported in American and British populations, but this is the first Australian study published[17,18,22]. A further strength of this study is the high power of the analyses: the comparison of the WOMAC domains in both surgical groups had a power well over 80%. A similar power was calculated for physical function and bodily pain in the SF-36 questionnaire. Adjustment for age, sex and disease duration was performed in multiple regression analysis.

In summary, WOMAC and MOS SF-36 detect significant and clinically meaningful changes in outcome after hip and knee replacement. We conclude that the WOMAC is a more responsive measure than the MOS SF-36, thus requiring a smaller sample size. This is important when designing short-term studies saving resources in time and money. In studies with a follow-up longer than 6 months MOS SF-36 provides useful additional information compared with using the WOMAC alone. For long-term studies it is useful to apply both the WOMAC and the MOS SF-36 questionnaires. After adjusting for age, gender and duration of OA the WOMAC outcomes of pain, stiffness and function and the SF-36 outcomes of pain, physical function, vitality and mental health following hip joint replacement surgery were greater than the improvement following knee joint replacement.

Acknowledgments

The study was funded by the Australian National Health and Medical Research Council. Dr C. Bachmeier was supported by the Ciba-Geigy-Jubiläumsstiftung, Karger Fonds, Schweizerische Gesellschaft für Rheumatologie and Schweizerische Akademie der Wissenschaften in Basel, Switzerland.

References

1. Lavernia CJ, Drakeford MK, Tsao AK, Gittelsohn A, Krackow KA, Hungerford DS. Revision and primary hip and knee arthroplasty. A cost analysis. Clin Orthop 1995;311:136–41.
2. Brooks P, Lapsley H, Oh ES, Cross M, Courtenay B, Schwarz J. *et al*. The economic impact of joint replacement surgery: a preliminary analysis. (Abstracts) Proceedings of APLAR Congress, April 21–26 1996. Melbourne: Australia.
3. Bellamy N. Osteoarthritis—an evaluative index for clinical trials. M.Sc. Thesis. 1982. McMaster University, Hamilton, Ontario, Canada.
4. 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 following total hip and knee arthroplasty in osteoarthritis. J Orthop Rheumatol 1988;1:95–108.
5. McHorney CA, Ware JE Jr, Rogers W, Raczec AE, Lu JFR. The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts: results from the medical outcome study. Med Care 1992;30(Suppl. 5):MS253–65.
6. McHorney CA, Ware JE Jr, Raczec AE. The MOS 36-item Short-Form Survey (SF-36). II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31: 247–63.
7. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al*. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24: 799–802.
8. Hochberg MC, Altman RD, Brandt KD, Moskowitz RW. Design and conduct of clinical trials in osteoarthritis: preliminary recommendations from a task force of the Osteoarthritis Research Society. J Rheumatol 1997;24:792–4.
9. SPSS. Release 6.0 (computer program). Chicago: SPSS Inc. 1993.
10. Bellamy N. WOMAC Osteoarthritis Index. A User's Guide. London, Ontario: N. Bellamy, 1995.
11. Ware JE, Snow KK, Kosinski M, Sandek B. SF-36 Health Survey Manual and Interpretation Guide, 1993. Boston, MA: The Health Institute, New England Medical Centre.
12. Bellamy N. Musculoskeletal clinical metrology. 1st ed. Dordrecht, Boston, London: Kluwer Academic Publishers, 1993.
13. Liang MH, Fossel AH, Larson MG. Comparison of five health status instruments for orthopedic evaluation. Med Care 1990;28:632–42.
14. Ritter MA, Albohm MJ, Keating EM, Farris PM, Meding JB. Comparative outcomes of total joint arthroplasty. J Arthroplasty 1995;10:737–41.
15. McGuignan FX, Hozack WJ, Moriarty L, Eng K, Rothman RH. Predicting quality of life outcomes

following total joint arthroplasty. J Arthoplasty 1995;10:742–7.
16. Wilcock GK. Benefits of total hip replacement to older patients and the community. BMJ 1978;2:37–9.
17. Kirwan JR, Currey HLF, Freeman MAR, Snow S, Young PJ. Overall long-term impact of total hip and knee joint replacement surgery on patients with osteoarthritis and rheumatoid arthritis. Br J Rheumatol 1994;33:357–60.
18. Kiebzak GM, Vain PA, Gregory AM, Mokris JG, Mauerhan DR. SF-36 general health status survey to determine patient satisfaction at short-term follow-up after total hip and knee arthroplasty. J South Orthop Assoc 1997;6:169–72.
19. Drewett RF, Minns RJ, Sibly TF. Measuring outcome of total knee replacement using quality of life indices. Ann Royal College of Surgeons England, 1992;74:286–90.
20. Williams A. Economics of coronary artery bypassing. BMJ 1985;291:326–9.
21. Stucki G, Liang MH, Phillips C, Katz JN. The Short Form 36 is preferable to the SIP as a generic health status measure in patients undergoing elective total hip arthroplasty. Arthritis Care and Research 1995;8:174–81.
22. Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) instrument in the measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22:1193–6.
23. Petrie K, Chamberlain K, Azariah R. The psychological impact of hip arthroplasty. Aust NZ J Surg 1994;64:115–17.