Quality of Life Research (2005) 14: 1141–1155
DOI 10.1007/s11136-004-4599-9

The impact of non-traumatic hip and knee disorders on health-related quality of life as measured with the SF-36 or SF-12. A systematic review

Johanna M. van der Waal[1], Caroline B. Terwee[1], Daniëlle A.W.M. van der Windt[1,2], Lex M. Bouter[1] & Joost Dekker[1,3]
[1]*Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, The Netherlands (E-mail: cb.terwee@vumc.nl);* [2]*Department of General Practice, VU University Medical Center, Amsterdam, The Netherlands;* [3]*Department of Rehabilitation Medicine, VU University Medical Center, Amsterdam, The Netherlands*

Accepted in revised form 11 October 2004

Abstract

Objectives: The purpose of this review is to summarize the available evidence on the impact of non-traumatic hip or knee disorders on health-related quality of life (HRQL), as measured with the Short Form 36 Health Survey (SF-36) or Short Form 12 Health Survey (SF-12), by comparing this with data from reference populations. *Methods*: Studies were identified by an electronic search of the MEDLINE, PsychInfo and Cinahl databases. Studies with the following features were included: study population included patients with non-traumatic hip or knee disorders, the SF-36 or SF-12 was used as an outcome measure and mean scores on these HRQL measures were presented. Using mean HRQL scores from the selected studies and scores from reference populations, *z*-scores were computed. Pooled estimates were computed for subgroups of studies with similar patients in similar settings. *Results*: A total of 40 studies met the inclusion criteria. Patients with non-traumatic hip and knee disorders scored up to 2.5 standard deviations (SDs) below reference population values, especially on the physical aspects of HRQL. Social and mental aspects were up to 1 SD below reference population values, especially in patients in clinical settings. *Conclusions*: The impact of non-traumatic hip or knee disorders on HRQL is substantial, especially on the physical aspects of HRQL.

Key words: Hip, Knee, Musculoskeletal diseases, Quality of life, Review, Systematic

Introduction

Hip and knee disorders are a frequent health problem. Estimates of the prevalence of osteoarthritis (OA) depend on variations in definition, but OA is thought to affect more than 10–12% of the population in the United States (US) [1]. The 12-months period prevalence of hip and knee disorders among adults in the general population in the Netherlands is estimated at 28% [2]. The prevalence increases with age [2]. In the Netherlands, every year 10 per 1000 persons visit their general practitioner (GP) with a new episode of hip complaints and 31 per 1000 persons with a new episode of knee complaints [3].

Hip and knee disorders have substantial consequences for public health, because of their strong impact on functional disability, health care costs, sick leave and work disability [4–7]. The United Nations, the World Health Organization (WHO), governments, professional and patients' organizations have therefore declared 2000–2010 the Bone and Joint Decade, with the aim of determining the burden of musculoskeletal diseases and improving the health-related quality of life of people with musculoskeletal conditions. Quantifying the health

burden of disorders is critical to decisions involving the allocation of limited health care resources.

The burden of a disease relates not only to its incidence and prevalence, but also to the impact of the disease on the (health-related) quality of life (HRQL) of the patients who suffer from it [8–10]. It is generally agreed upon that HRQL encompasses several different dimensions including physical, emotional and social functioning [11]. To facilitate interpretation of HRQL data and to put population scores into perspective, generic instruments, such as the Short Form 36 Health Survey (SF-36), are usually best suited: generic measures facilitate comparisons of scores with those from a reference population, and enable a comparison of HRQL across different disease groups.

Although HRQL in patients with hip and knee disorders has been studied extensively using generic instruments, a systematic review about the impact of hip and knee disorders, such as OA, rheumatoid arthritis, bursitis, or non-specific hip or knee pain, on HRQL is not yet available.

By pooling the results of separate studies, a more precise estimate of the impact on HRQL can be made. The purpose of this review was to summarize the available evidence on the impact of non-traumatic hip or knee disorders on HRQL. In order to facilitate interpretation of the results and to quantify the impact of hip and knee complaints on HRQL, we wanted to compare patient scores with reference data from the general population. An extensive search of the literature showed that the SF-36 and SF-12 were the only instruments for which reliable country-specific reference data were available. Therefore, we only included studies that used the SF-36 or SF-12 to assess HRQL in our systematic review.

Methods

Search strategy and study selection

The SF-36 [12] and SF-12 [13] are the most commonly used generic HRQL measures in patients with hip or knee disorders. We conducted a systematic literature search to identify studies measuring HRQL, using the SF-36 or SF-12, in patients with non-traumatic hip or knee disorders. Publications were selected from the following databases: Medline (1966 until January 2003), PsycINFO (1977 until January 2003) and Cinahl (1982 until January 2003). No language restrictions were imposed.

The search terms used were (*SF-36* or *SF-12* or *SF-36* and *SF-12* or '*short form-36*' or '*short form-12*' or '*Short Form 36 Health Survey*' or '*Short Form thirty-six health survey*' or '*Short Form 12 Health Survey*' or '*Short Form twelve health survey*') plus *hip* or *knee* (MESH or free text word). The following criteria were used for inclusion of studies in the review:

- The study population consists of patients with non-traumatic hip or knee disorders.
- HRQL, defined as physical, emotional or social functioning, is measured using the SF-36 or SF-12.
- The mean scores of patients with non-traumatic hip or knee disorders on the SF-36 or SF-12 are presented in the article.

The following criteria were used for exclusion of studies from the review:

- The study population consists of patients after surgery. Our aim was to quantify the impact of a disorder and not the impact of a treatment.
- The study population consists of patients with traumatic injuries.
- The study population contains fewer than 100 patients with non-traumatic hip or knee disorders. This is an arbitrary cut-off point, but a sample of at least 100 persons will provide more reliable estimates of HRQL.
- The study population consists of patients participating in a (randomized) clinical trial (RCT). Due to strict selection criteria these patients are often not representative for members of the general population or patients encountered in everyday primary or secondary care.

The first author screened all titles, in order to exclude articles that obviously did not meet the selection criteria. Two independent reviewers (DW and JW) scanned all remaining abstracts. During a consensus meeting all abstracts that appeared to meet the selection criteria were selected. The full text of these articles was retrieved to select articles that met all selection criteria. When multiple articles used the same data, we included the most recent or most complete article. Finally, we

handsearched the reference lists of all included articles to find additional eligible studies.

Data extraction

We extracted data from each article on the following study characteristics: mean age of the study population, study size, country, setting, and case definition (diagnosis, disease stage). Extracted mean HRQL scores were entered into a custom-made spreadsheet.

To examine the impact of hip and knee disorders on HRQL we compared the scores on the SF-36 or SF-12 with scores obtained from country-specific reference populations [14–23]. These reference populations consist of a representative sample from the general population and are usually used as normative data. If possible, age- and sex-specific reference data were used. Details about which reference populations were used are provided in the Results section.

Data analysis

We computed *z*-scores for each subscale of the SF-36 or SF-12 by dividing the difference between the mean HRQL score of the patient group and the mean HRQL score of the reference population by the standard deviation of the mean HRQL score of the reference population. We computed *z*-scores for the SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) in the same way. The use of *z*-scores ('norm-based scoring') has been recommended by the developers of the SF-36 [24]. To compute *z*-scores, we used country-specific normative data whenever available. Tables 1–3 show which data were used to compute *z*-scores for each individual study.

No reference data from Switzerland were available. Instead, we used data from a French population [15]. Since France is a neighbouring country and a substantial number of inhabitants speak the same language, we thought these reference data were most suitable. When a study was conducted in multiple countries, reference data were used from the country that represented most patients. Lingard et al. [25] performed a multicentre study in the UK (six centres), the US (four centres) and Australia (two centres). Most centres were in the UK so reference data from an English population

Table 1. Studies in the general population, among community-living elderly or in primary care

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Birrell et al., 2000 [64]	195	63 (11)	UK	General practice and primary care rheumatology	Patients presenting with new episodes of hip pain in primary care	Representative sample of population aged 16 and over living in private households in the UK [14]
Carmona et al., 2001 [61]	223	$\geqslant$ 20	Spain	General population	Symptomatic knee OA, ACR classification criteria	Representative sample of Spanish population [22]
Clark et al., 1998 [27]	415 men	Median 66, 22–90	US	Male outpatients at Veterans Affairs medical centres	Knee OA according to three-questions	Representative sample of US population [22]
Cooper and Kohlmann, 2001 [62]	132,514	74 (6.1), 65–108	US	Medicare beneficiaries, aged 65 and older; community living elderly	Arthritis of hip or knee, diagnosis based on symptoms, and co-morbidity	Representative sample of US population [22]
Wilcox et al., 2000 [63]	429	72 (5.0), 65–88	US	Elderly enrolled in the Observational Arthritis Study in Seniors (community-based study)	Knee pain, or knee pain with radiographic evidence of OA	US general population norms [17]

SD: standard deviation, UK: United Kingdom, US: United States, OA: osteoarthritis, ACR: American College of Rheumatology.

Table 2. Studies in outpatient clinics and other clinical settings

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Angst et al., 2001 [30]	211	65 (10.0), 37–86	Switzerland	Rheumatology and rehabilitation clinic	Patients with hip or knee OA referred for a comprehensive inpatient rehabilitation programme	Representative sample of French population [15, 22]
Brazier et al., 1999 [33]	112	64	UK	Rheumatology outpatient clinic	All new patients with diagnosis of knee OA according to hospital rheumatologist or orthopaedic specialist	Representative sample of population aged 16 and over living in private households in the UK [14]
Ren et al., 1998 [31]	328	?*	US	Veterans Health Study, four Veterans Affairs outpatient clinics	Patients with OA, based on patient's report of having a physician's diagnosis, treatment, and/or symptoms	US general population norms [17]
Thumboo et al., 2002 [32]	110	61, 33–86	Singapore	Tertiary hospital	Inpatients or outpatients with knee OA	Random sample, district of Singapore, including both public and private housing projects, age between 21 and 65 [16]
Wolfe et al., 2000 [29]	648	68 (11.7)	US	Departments of rheumatology	Patients with OA of the hip or knee, first outpatient visit	Representative sample of US population [22]

SD: standard deviation, UK: United Kingdom, US: United States, OA: osteoarthritis, * ?: not presented in the article.

were used [14]. Mahomed et al. [26] performed a study in the US and Canada. As most patients were from the US (60%), the American reference data for the SF-36 [17] and the SF-12 [22] were used.

Clark et al. [27] used the SF-36 PCS and MCS to assess HRQL among male outpatients in the US. As no sex-specific reference data are available for SF-36 PCS and MCS scores, we used the available reference data from a representative sample of the American population [22]. Spanish reference data for the SF-36 were sex-specific. For the study by Escobar et al. [28] we decided to use female reference data from Spain [23], because most patients in this Spanish study were female (TKA 56% and THA 71%).

Data synthesis and presentation

Results are presented separately for hip and knee disorders (if possible) and for the following study settings: general population, community-dwelling elderly, primary care, clinic (hip or knee OA), and in patients prior to total hip arthroplasty (THA) or total knee arthroplasty (TKA). For studies concerning similar patient populations (with respect to diagnosis or mean age) and similar settings, pooled estimates were computed for separate dimensions of SF-36 or SF-12, weighted by study size. Studies among community-living elderly showed many differences regarding characteristics of the patient population (case definition, setting or age). Consequently, we refrained from computing pooled *z*-scores for these studies. We pooled two studies concerning patients with knee or hip OA in clinical settings (data from hip and knee patients were not presented separately in these studies) [29, 30], three studies on patients with knee OA in clinical settings [31–33], two studies on patients with knee or hip OA, admitted or scheduled for TKA or THA (data from hip and knee patients were also not presented separately in these studies) [34, 35], 19 studies on patients with knee OA admitted or scheduled for TKA [25, 26, 28, 33, 36–50] and 23 studies on patients with hip OA admitted or scheduled for THA [26, 28, 36, 39–46, 48–59]. Deviations of more than 0.5 Standard deviation (SD) from the reference population (z-score < -0.5) were considered clinically

Table 3. Studies in patients prior to TKA or THA

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Arslanian and Bond, 1999 [50]	TKA 949 THA 570	TKA 69, 36–93 THA 68, 20–89	US	?*	Patients scheduled for TKA or THA	US general population norms [17]
Bachmeier et al., 2001 [36]	TKA 108 THA 86	TKA 72 (7.0) THA 65 (11.5)	Australia	Four hospitals in Sydney	Patients with OA of the hip or knee, admitted for TKA or THA	Australian population norms [19]
Bayley et al., 1995 [42]	TKA 117 THA 90	?	US	Orthopaedic surgery clinics	Patients scheduled for TKA or THA	US general population norms [17]
Benroth and Gawande, 1999 [45]	TKA 110 THA 63	TKA 66, 30–88 THA 66, 30–88	US	?	Consecutive patients scheduled for TKA or THA	US general population norms [17]
Brazier et al., 1999 [33]	TKA 118	TKA 71, 47–87	UK	Rheumatology clinic	Patients with knee OA according to hospital rheumatologist or orthopaedic specialist, admitted for TKA	Representative sample of population aged 16 and over living in private households in the UK [14]
Croft et al., 2002 [57]	THA 611	THA median: 71	UK	Secondary care orthopaedic centre	Patients with hip OA on waiting list for THA	Representative sample of population aged 16 and over living in private households in the UK [14]
Dawson et al., 1996 [58]	THA 219	THA median: 71 (11), 36–90	UK	Preadmission assessment clinic, Nuffield Orthopaedic Centre, Oxford	Consecutive patients referred for THA	Representative sample of population aged 16 and over living in private households in the UK [14]
Dawson et al., 1996 [59]	THA 173	THA median: 71, 38–89	UK	Preadmission assessment clinic	Patients with OA or RA scheduled for unilateral THA	Representative sample of adult population aged 16 and over living in private households in the UK [14]
Dervin et al., 2003 [47]	TKA 126	TKA 62 (8.6)	Canada	Orthopaedic outpatient clinic, Ottawa General Hospital	All patients with OA of the knee, referred for TKA	Canadian normative data [18]

Table 3. Continued

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Escobar et al., 2002 [28]	TKA 100 THA 103	TKA 71 (6) THA 69 (10.3)	Spain	Tertiary hospital	Patients with hip or knee OA on waiting list for TKA or THA	Randomly selected non-institutionalized female population 65 years and over in Spain [23]
Fortin et al., 1999 [48]	TKA 106 THA 116	TKA 67 THA 67	Canada	Brigham and Women's Hospital in Boston and Montreal General Hospital	All patients with hip or knee OA scheduled for elective primary TKA or THA	Canadian normative data [18]
Fortin et al., 2002 [49]	TKA 81 THA 84	TKA 68 (9.1) THA 66 (8.2)	Canada	Boston and Montreal teaching hospitals	Consecutive patients with OA scheduled for elective TKA or THA	Canadian normative data [18]
Hartley et al., 2002 [37]	TKA 100	TKA 76, 47–90	UK	?	Consecutive patients scheduled for primary TKA	Representative sample of the UK population [22]
Hashimoto et al., 2003 [65]	TKA 428 TKA 261 TKA 170 TKA 113	71 (9) 69 (10) 69 (8) 68 (10)	UK US Australia Canada	International prospective cohort	Patients scheduled for primary TKA	Representative sample of the UK population [22], US [17], Australian [19] and Canadian [18] normative data
Heck et al., 1998 [38]	TKA 291	TKA 70, 50–88	US	Orthopaedic surgery clinics, State of Indiana	Patients with knee OA in community practice, referred for TKA	US general population norms [17] and representative sample of US population [22]
Hozack et al., 1997 [39]	TKA 149 THA 151	TKA 70, 40–100 THA 64, 28—86	US	?	Patients undergoing surgery for degenerative arthritis of the knee or hip	US general population norms [17]
Jones et al., 2001 [40]	TKA 256 THA 197	TKA 71 THA 71	Canada	Community-based cohort, health care region	Patients recommended for primary TKA or THA	Canadian normative data [18]
Jones et al., 2000 [41]	TKA 276 THA 228	TKA 68 (10.1) THA 68 (10.1)	Canada	Community-based cohort, health care region	Patients recommended for primary TKA or THA, at least 7 days before surgery	Canadian normative data [18]

Table 3. Continued

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Kelly et al., 2001 [34]	313	68, 27–89	Canada	Two referral hospitals, departments of orthopaedic surgery	Consecutive patients recommended for TKA or THA	Canadian normative data [18]
Kiebzak et al., 2002 [43]	TKA 415 THA 207	Patients of all ages were included	US	Orthopaedic clinic	Patients scheduled for elective primary TKA or THA	US general population norms [17]
Kiebzak et al., 1997 [44]	TKA 78 THA 80	TKA: men 59(15), women 68(12) THA: men 66 (8) Women 68 (10)	US	Miller Orthopaedic Clinic	Consecutive patients scheduled for TKA or THA	US general population norms [17]
Lingard et al., 2001 [25]	TKA 697	TKA 70, 38–90	UK, US, Australia	Kinemax Outcomes Study; 12 centres in US, UK, and Australia	Patients with primary diagnosis of OA and no history of knee implant surgery, scheduled for TKA	Representative sample of population aged 16 and over living in private households in the UK [14]
Mahomed et al., 2002 [26]	TKA 89 THA 103	TKA 68(9) THA 66(9)	US, Canada	Two tertiary referral centres in Boston and Montreal	Patients undergoing primary TKA or THA	US general population norms [17] and representative sample of US population [22]
Mangione et al., 1997 [51]	THA 236	THA 67(9)	US	University tertiary care hospital	Patients admitted for THA	US general population norms [17]
Nilsdotter et al., 2001 [52]	THA 160	THA 61, 50–72	Sweden	Dept of orthopaedics, Halmstad	Patients with primary OA scheduled for THA	Normative data from seven general population studies in Sweden [20]
Nilsdotter and Lohmander, 2002 [53]	THA 124	THA 71, 51–88	Sweden	Dept of Orthopaedics, Halmstad	Consecutive patients with primary OA scheduled for THA	Normative data from seven general population studies in Sweden [20]
O'Connell et al., 2000 [54]	THA 100	THA > 50	Ireland	North Eastern Health Board Regional Orthopaedic Unit, Navan	Consecutive patients undergoing primary THA	Random sample from electoral register in Ireland [21]
O'Shea et al., 2002 [55]	THA 144	?	Ireland	Cappagh National Orthopaedic Hospital	Information from on-going in-hospital audit of THA	Random sample from electoral register in Ireland [21]
Salmon et al., 2001 [46]	TKA 53 THA 107	TKA 66(11.1) THA 69(11.0)	UK	Two teaching hospitals	Patients admitted for unilateral, primary elective TKA or THA	Representative sample of population aged 16 and over living in private households in the UK [14]

Table 3. Continued

Author, year [reference]	*n*	Age mean (SD), range	Country	Setting	Case definition	Reference data
Singer et al., 1999 [56]	THA 284	THA 70(10.3)	Canada	?	Patients with hip OA on waiting list for THA	Canadian normative data [18]
Williams et al., 1997 [35]	209	TKA 69 THA 62	Canada	Teaching or community hospitals in Ontario, departments of orthopaedic surgery	Patients referred for TKA or THA	Canadian normative data [18]

SD: standard deviation, UK: United Kingdom, US: United States, OA: osteoarthritis, TKA: total knee arthroplasty, THA: total hip arthroplasty, * ?: not presented in the article.

important, based on recent guidelines for HRQL research [60].

Results

Search

The results of our search strategy are presented in Figure 1. The search strategy identified 323 abstracts. After initial selection by the first author 126 abstracts were excluded. Two independent reviewers assessed the remaining 197 abstracts. Another 135 abstracts were excluded. For the remaining 62 abstracts the full text article was retrieved. Twenty-two of these 62 articles were excluded. Forty articles contained mean scores from the SF-36 or SF-12, and were included in the review. Handsearching the reference lists of these 40 articles did not result in the identification additional relevant articles.

Characteristics of the study populations

The studies contained data about the HRQL of patients with hip and knee disorders in five different settings. Of the 40 studies we included, one study concerned patients in the general population [61] (223 patients), three studies concerned community living elderly [27, 62, 63] (in total 133,358 patients), one study concerned patients in a primary care setting [64] (195 patients), five studies concerned patients in clinical settings [29–33] (in total 1407 patients) and thirty studies concerned patients with OA admitted or scheduled for total arthroplasty (in total 5191 patients admitted for TKA [25, 26, 28, 33, 36–50, 65] and 4236 patients admitted for THA [26, 28, 36, 39–46, 48–59]) The characteristics of these studies are described in Tables 1–3. One article [33] describes HRQL of patients in clinical settings and of patients prior to TKA and therefore appears in Tables 2 and 3.

The study in the general population concerned patients with symptomatic knee OA. The studies concerning community-living elderly included patients with knee pain, or knee or hip OA. The study in primary care concerned patients presenting with new episodes of hip pain. The studies in outpatient clinics, rheumatology and rehabilitation clinics or

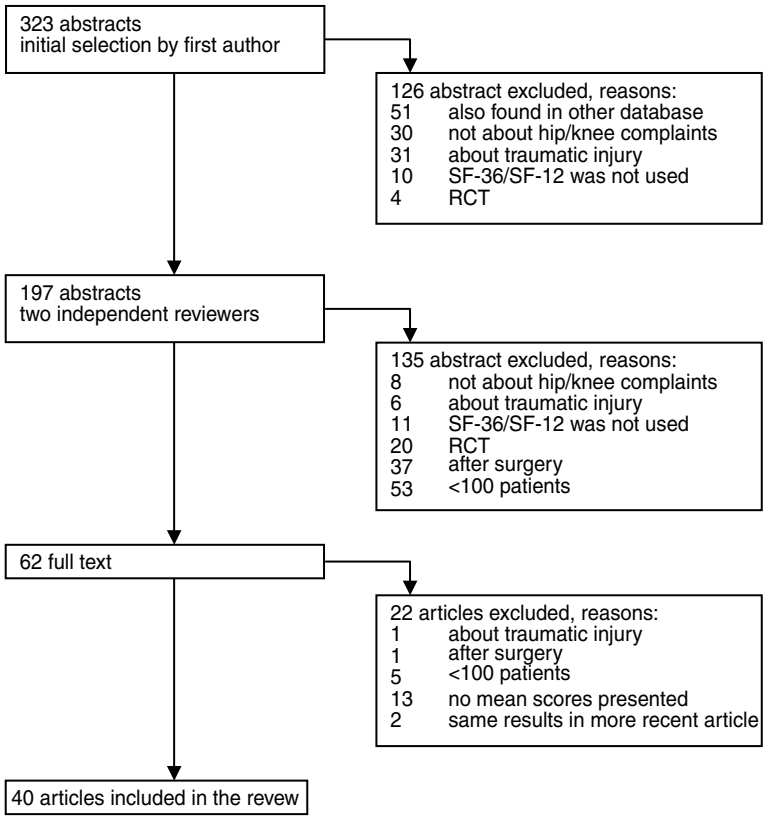

Figure 1. Study selection.

tertiary hospitals, all included patients with hip or knee OA or patients prior to TKA or THA.

HRQL in patients with hip or knee disorders

In Table 4 (pooled) *z*-scores are presented to show the impact of hip and knee complaints on HRQL. Table 4 demonstrates, for example, that elderly patients with knee OA in the general population in the US [63] scored on average 0.3 SDs below the reference population on the SF-36 general health subscale. Patients prior to THA (pooled estimate of 23 studies) scored on average 2.5 SDs below the reference population on the SF-36 physical functioning subscale, and 2.0 SDs below the reference population on the SF-36 role limitations in physical functioning subscale.

In studies using the PCS and MCS summary scores, patients scored approximately 1.5–2.5 SDs below the reference population for the SF-36 PCS subscale and up to 0.5 SD below the reference population for the SF-36 MCS subscale. Scores on the SF-36 PCS and MCS scales were about 1 SD lower in patients admitted for surgery than in community living elderly.

Two studies used the SF-12. Carmona et al. [61] performed a study in the general population in Spain, and Hartley et al. [37] assessed HRQL in patients scheduled for TKA in the UK. The results show poorer scores for patients in the Spanish general population, compared to patients scheduled for TKA in the UK.

In Figure 2, the data from Table 4 are presented graphically. Figure 2 demonstrates that the profile for the different subscales shows a similar pattern of reduction in HRQL. Patients from all settings scored approximately 1–2 SDs below the reference population for three subscales of the SF-36: physical functioning, role limitations in physical functioning, and bodily pain. The scores on mental health in patients in the general population, primary care or among community-living elderly were rather

Table 4. Impact on health-related quality of life: *z*-scores on dimensions of SF-36 and SF-12: pooled *z*-scores are presented for subgroups of studies with similar patients in similar settings

Patient sample	Study	PF	RP	BP	GH	VT	SF	RE	MH	PCS	MCS	SF-12 Ph	SF-12 Ps
Gen. pop., knee OA, Spain	Carmona et al., 2001 [61]											−2.0	−0.9
Comm. living elderly, OA hip or knee, US	Cooper and kohlmann, 2001 [62]									−1.5	0.2		
Gen. pop., elderly, knee OA, US	Wilcox et al., 2000 [63]				−0.3								
Primary care, hip pain, UK	Birrell et al., 2000 [64]	−1.0	−1.6	−1.2	−0.2	−0.4	−0.4	0.5	0.1				
Primary care, knee OA, US, male	Clark et al., 1998 [27]									−2.1	−0.2		
Clinic, knee or hip OA	Pooled *z*−scores [29, 30] (Range)	−2.4	−1.7	−1.8	−0.7	−0.9	−0.6	−0.8	−0.2	−2.4 (−2.84 to −2.29	−0.5 (−0.62 to 0.0)		
Clinic, knee OA	Pooled *z*−scores [31–33] (Range)	−1.8 (−2.00 to −1.72)	−1.6 (−1.91 to −1.59)	−1.3 (−1.66 to −1.08)	−1.1	−1.3	−1.3	−1.4	−0.7				
Prior to TKA/THA, Ca	Pooled *z*−scores [34, 35] (Range)	−3.1 (−3.18 to −2.87)	−1.9 (−2.08 to −1.70)	−1.9 (−2.01 to −1.80)	−0.8 (−0.97 to −0.45)	−1.2 (−1.29 to −1.07)	−1.5 (−1.58 to −1.26)	−0.7 (−1.14 to −0.16)	−0.4 (−0.56 to −0.06)				
Prior to TKA	Pooled *z*−scores [25, 26, 28, 33, 36–50, 65] (Range)	−2.4 (−3.27 to −1.05)	−1.9 (−2.13 to −1.66)	−1.6 (−1.95 to −0.76)	−0.4 (−0.86 to 0.00)	−0.8 (−1.35 to −0.48)	−0.9 (−1.63 to −0.54)	−0.7 (−1.40 to −0.43)	−0.1 (−0.56 to −0.07)	−2.5 (−2.74 to −1.98)	0.0 (−0.19 to 0.61)	−1.8	0.0
Prior to THA	Pooled *z*−scores [26, 28, 36, 39–46, 48–59] (Range)	−2.5 (−3.35 to −1.23)	−2.0 (−2.36 to −1.70)	−1.8 (−2.16 to −0.88)	−0.3 (−0.90 to 0.58)	−0.8 (−1.49 to −0.13)	−1.3 (−1.93 to −0.59)	−0.8 (−1.77 to −0.19)	−0.3 (−0.75 to 0.32)	−2.8 (−2.94 to −2.43)	−0.2 (−0.41 to 0.61)		

PF: physical functioning, RP: role limitations in physical functioning, BP: bodily pain, GH: general health, VT: vitality, SF: social functioning, RE: role limitations in emotional functioning, MH: mental health, PCS: physical component score, MCS: mental component score, SF-12 Ph: SF-12 physical subscale, SF-12 Ps: SF-12 psychological subscale, OA: osteoarthritis, US: United States, UK: United Kingdom, Ca: Canada, TKA: total knee arthroplasty, THA: total hip arthroplasty Data in brackets represent the range of the individual studies that were pooled. Empty cells: the article does not report that subscale.

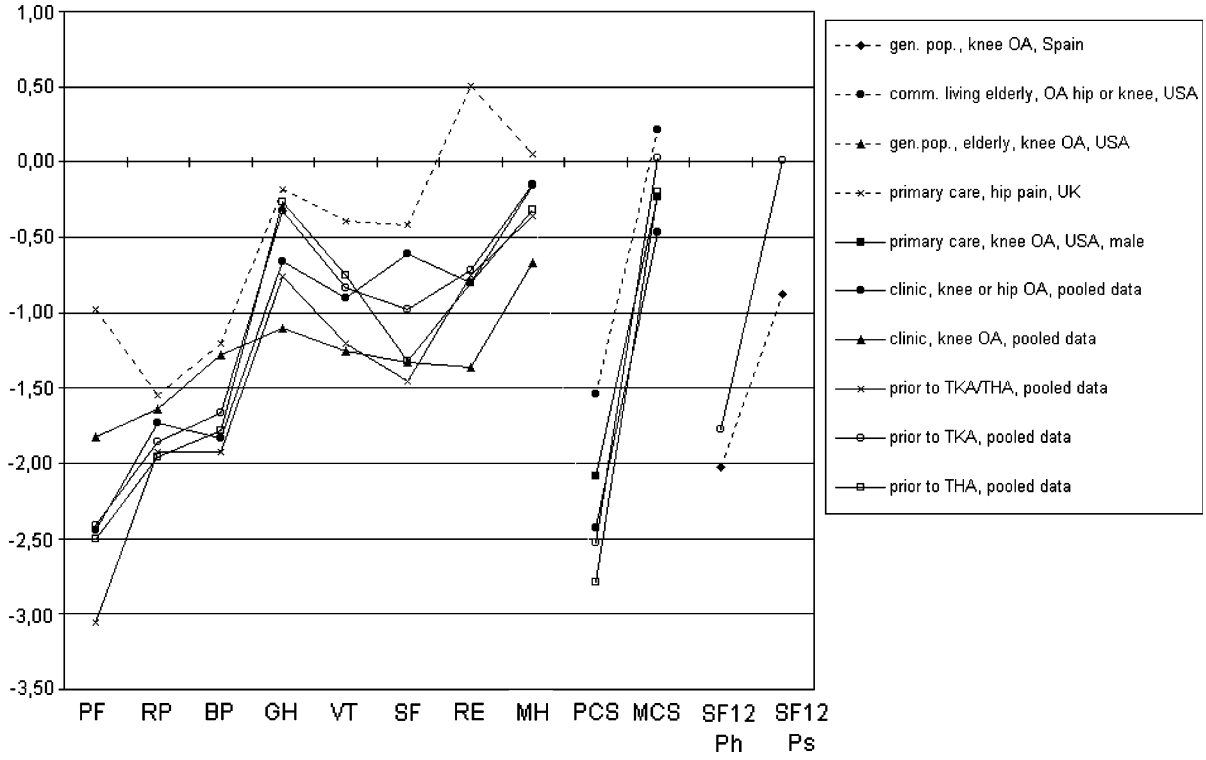

Figure 2. *z*-Scores for each subscale of the SF-36, the two summary scales of the SF-36, and the SF-12 subscales in different patient populations. PF: physical functioning, RP: role limitations in physical functioning, BP: bodily pain, GH: general health, VT: vitality, SF: social functioning, RE: role limitations in emotional functioning, MH: mental health, PCS: physical component score, MCS: mental component score, SF-12 ph: SF-12 physical subscale, SF-12 Ps: SF-12 psychological subscale

similar to scores from reference populations. Scores on mental health were worse in patients in clinical settings. Clinical outpatients and patients admitted for surgery (TKA or THA) scored up to1 SD below the reference population for mental health.

Discussion

The results of this systematic review of 40 observational studies measuring HRQL in patients with non-traumatic hip or knee disorders show that these disorders have a substantial impact on HRQL. According to previous studies, the threshold for clinically important changes in health-related quality of life appears to be half a SD [60]. The scores on subscales with physical components were especially low: up to 2 SDs below reference values. Scores on subscales with mental and social components were only low for patients seen in clinical settings and those admitted for THA or TKA (up to 1 SD below reference values). These results indicate that patients from all settings were markedly impaired in their physical functioning, and that patients in clinical settings (who may have more severe hip or knee disorders) generally have a poor HRQL. Remarkably, the profile among the various subscales is about the same among all patient groups.

To put the results of our review into perspective, we compared the scores from the patient groups included in our review with those obtained from patients with other disorders, retrieved from the literature. This comparison shows that patients with hip or knee disorders have poorer scores on several dimensions of HRQL than patients with heart disease or cancer [66]. Patients with heart disease or cancer score up to 1.2 SD below reference values on physical subscales [67]. Although some of the HRQL effects seen in OA patients may have been caused by comorbidities, such as hypertension or cardiovascular disease [68], that affect persons

with OA, this comparison suggests that hip or knee disorders can have a substantial effect on HRQL.

The results of this review show that patients with hip or knee disorders generally have high pain scores, which limits physical and social functioning. This draws attention to the importance of pain management and coping with pain in these patient groups. Interventions should be developed and evaluated that are directed towards reduction of pain, improvement of functional capacity, and HRQL of patients with hip or knee problems.

The population in western countries is aging, and an increasing number of people are suffering from hip or knee complaints. Most of these patients are encountered and cared for in primary care. However, only one study measured HRQL in patients with hip or knee disorders in a primary care setting. This study [64] concerned an older population, which may limit the possibilities for generalizing these results to other primary care populations. More research should be aimed at assessing and improving HRQL in patients with hip or knee problems in primary care.

Although the SF-36 is widely available and validated in many languages, for most reference populations, age- and sex-specific data are not available, and for one study included in this review data from a country-specific reference population were not available at all [30]. Furthermore, one study was performed in male patients and no sex-specific reference data were available [27]. Another study was performed in both male and female patients, but only sex-specific reference data were available [28]. These factors mean that our *z*-scores may have been somewhat biased because we could not use an appropriate reference population. Women generally score poorer on HRQL questionnaires than men. Thus, in studies for which we had to use female reference data, we may have slightly underestimated *z*-scores. In addition, the mean age of the reference population was in some studies lower than the mean age of the study population. This means that the burden of hip and knee disorders on the physical subscales may have been slightly overestimated. In the US reference population, the scores on the physical subscales for people aged 65–74 years were about 0.2–0.5 SD lower than the mean scores of the general population [17]. This may give an indication of the amount of overestimation in this study due to younger age of a reference population.

To our knowledge guidelines for pooling *z*-scores are not yet available. We decided to compute pooled *z*-scores weighted by study size. The inverse of the variance of an estimate is more often used as a weighting factor in meta-analysis. However, in our review the variance is part of the outcome of each study (the *z*-score). Using the inverse of the variance as a weighting factor would mean that studies that show a wide range in HRQL scores (which may accurately reflect HRQL in the assessed population) would receive less weight in the pooled estimate. This would have affected the results of our review somewhat.

We only selected studies using the SF-36 or SF-12 for our review. Other generic and disease-specific measures are available and have been used in studies concerning patients with non-traumatic hip and knee complaints. For these measures often no information is available as to what scores represent important limitations in health, and data from reference populations are generally not available, which hampers the interpretation of absolute scores. More attention should be paid to the meaning of absolute scores of health status questionnaires, and age- and sex-specific reference data of the general population should be provided.

The study in the general population in Spain is said to consist of patients with knee OA, obtained from a representative sample of the Spanish general population [61]. Unexpectedly, these results showed a lower *z*-score for these patients than for patients admitted for TKA, which raises doubt about the representativeness of these patients with knee OA.

In conclusion, this is the first review to quantify the impact of non-traumatic hip or knee disorders on HRQL. This impact turns out to be substantial, especially the impact on physical aspects of HRQL. The results of this review support the effort of the organizers of the Bone and Joint Decade [10] to determine the burden of musculoskeletal diseases and underscore their statement that the HRQL of people with musculoskeletal conditions should be improved.

Acknowledgement

This study is financially supported by the Dutch Arthritis Association. Grant number: NR 99-2-401.

References

1. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini EH, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998; 41: 778–799.
2. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: Prevalences, consequences and risk groups, the DMC(3)-study. Pain 2003; 102: 167–178.
3. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose. Episodegegevens uit de huisartspraktijk [From complaint to diagnosis: episode data from family practice]. Bussum: Uitgeverij Coutinho, 1998.
4. Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. Lancet 1999; 354: 1248–1252.
5. Straaton KV, Fine PR. Addressing work disability through vocational rehabilitation services. Bull Rheum Dis 1997; 46: 1–3.
6. Hoffman C, Rice D, Sung HY. Persons with chronic conditions. Their prevalence and costs. JAMA 1996; 276: 1473–1479.
7. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Association of locomotor complaints and disability in the Rotterdam study. Ann Rheum Dis 1995; 54: 721–725.
8. Reginster JY, Khaltaev NG. Introduction and WHO perspective on the global burden of musculoskeletal conditions. Rheumatology (Oxford) 2002; 41(Suppl 1): 1–2.
9. Woolf AD, Akesson K. Understanding the burden of musculoskeletal conditions. The burden is huge and not reflected in national health priorities. BMJ 2001; 322: 1079–1080.
10. Focal Point of Bone and Joint Decade is Determining Burden of Disease. American College of Rheumatology News, 2001, p. 20.
11. Testa MA, Simonson DC. Assesment of quality-of-life outcomes. N Engl J Med 1996; 334: 835–840.
12. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473–483.
13. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998; 51: 1171–1178.
14. Erens B, Primatesta P, Prior G. Health Survey for England. London: The Stationery Office, 1996.
15. Leplege A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: Translation, cultural adaptation and preliminary psychometric evaluation. J Clin Epidemiol 1998; 51: 1013–1023.
16. Thumboo J, Chan SP, Machin D, Soh CH, Feng PH, Boey ML, et al. Measuring health-related quality of life in Singapore: Normal values for the English and Chinese SF-36 Health Survey. Ann Acad Med Singapore 2002; 31: 366–374.
17. Ware JE, Snow KK, Kosinski M, Gandek B. Health Survey Manual and Interpretation Guide. Boston, MA: Nimrod Press, 1993.
18. Hopman WM, Towheed T, Anastassiades T, Tenenhouse A, Poliquin S, Berger C, et al. Canadian normative data for the SF-36 health survey. Canadian Multicentre Osteoporosis Study Research Group. CMAJ 2000; 163: 265–271.
19. Australian Bureau of Statistics. National Health Survey Australia, 1995: SF-36 Population Norms (Catalogue no 4399.0). Canberra: Australian Bureau of Statistics, 1996.
20. Sullivan M, Karlsson J, Ware JE Jr. 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: 1349–1358.
21. Blake C, Codd MB, O'Meara YM. The Short Form 36 (SF-36) Health Survey: normative data for the Irish population. Ir J Med Sci 2000; 169: 195–200.
22. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. J Clin Epidemiol 1998; 51: 1171–1178.
23. Regidor E, Barrio G, de la FL, Domingo A, Rodriguez C, Alonso J. Association between educational level and health related quality of life in Spanish adults. J Epidemiol Community Health 1999; 53: 75–82.
24. Ware JE, Koskinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. 2nd ed. Boston, MA: The Health Institute, 1994.
25. Lingard EA, Katz JN, Wright RJ, Wright EA, Sledge CB. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. J Bone Joint Surg Am 2001; 83-A: 1856–1864.
26. Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J Rheumatol 2002; 29: 1273–1279.
27. Clark JA, Spiro A III, Fincke G, Miller DR, Kazis LE. Symptom severity of osteoarthritis of the knee: A patient-based measure developed in the veterans health study. J Gerontol A Biol Sci Med Sci 1998; 53: M351–M360.
28. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validation of the Spanish Version of the WOMAC Questionnaire for Patients with Hip or Knee Osteoarthritis. Clin Rheumatol 2002; 21: 466–471.
29. Wolfe F, Kong SX, Watson DJ. Gastrointestinal symptoms and health related quality of life in patients with arthritis. J Rheumatol 2000; 27: 1373–1378.
30. Angst F, Aeschlimann A, Steiner W, Stucki G. Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. Ann Rheum Dis 2001; 60: 834–840.
31. Ren XS, Kazis L, Lee A, Miller DR, Clark JA, Skinner K, et al. Comparing generic and disease-specific measures of physical and role functioning: Results from the Veterans Health Study. Med Care 1998; 36: 155–166.
32. Thumboo J, Chew LH, Lewin-Koh SC. Socioeconomic and psychosocial factors influence pain or physical function in Asian patients with knee or hip osteoarthritis. Ann Rheum Dis 2002; 61: 1017–1020.

33. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999; 38: 870–877.
34. Kelly KD, Voaklander DC, Johnston DW, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. J Arthroplasty 2001; 16: 351–359.
35. Williams JI, Llewellyn TH, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1997; 3: 59–68.
36. Bachmeier CJ, March LM, Cross MJ, Lapsley HM, Tribe KL, Courtenay BG, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthr Cartilage 2001; 9: 137–146.
37. Hartley RC, Barton-Hanson NG, Finley R, Parkinson RW. Early patient outcomes after primary and revision total knee arthroplasty. A prospective study. J Bone Joint Surg Br 2002; 84: 994–999.
38. Heck DA, Robinson RL, Partridge CM, Lubitz RM, Freund DA. Patient outcomes after knee replacement. Clin Orthop 1998; 93–110.
39. Hozack WJ, Rothman RH, Albert TJ, Balderston RA, Eng K. Relationship of total hip arthroplasty outcomes to other orthopaedic procedures. Clin Orthop 1997; 88–93.
40. Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. The effect of age on pain, function, and quality of life after total hip and knee arthroplasty. Arch Intern Med 2001; 161: 454–460.
41. Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. J Rheumatol 2000; 27: 1745–1752.
42. Bayley KB, London MR, Grunkemeier GL, Lansky DJ. Measuring the success of treatment in patient terms. Med Care 1995; 33: AS226–AS235.
43. Kiebzak GM, Campbell M, Mauerhan DR. The SF-36 general health status survey documents the burden of osteoarthritis and the benefits of total joint arthroplasty: But why should we use it? Am J Manage Care 2002; 8: 463–474.
44. 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: 169–172.
45. Benroth R, Gawande S. Patient-reported health status in total joint replacement. J Arthroplasty 1999; 14: 576–580.
46. Salmon P, Hall GM, Peerbhoy D, Shenkin A, Parker C. Recovery from hip and knee arthroplasty: Patients' perspective on pain, function, quality of life, and well-being up to 6 months postoperatively. Arch Phys Med Rehabil 2001; 82: 360–366.
47. Dervin GF, Stiell IG, Rody K, Grabowski J. Effect of arthroscopic debridement for osteoarthritis of the knee on health-related quality of life. J Bone Joint Surg Am 2003; 85-A: 10–19.
48. Fortin PR, Clarke AE, Joseph L, Liang MH, Tanzer M, Ferland D, et al. Outcomes of total hip and knee replacement: Preoperative functional status predicts outcomes at six months after surgery. Arthritis Rheum 1999; 42: 1722–1728.
49. Fortin PR, Penrod JR, Clarke AE, St Pierre Y, Joseph L, Belisle P, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46: 3327–3330.
50. Arslanian C, Bond M. Computer assisted outcomes research in orthopedics: Total joint replacement. J Med Syst 1999; 23: 239–247.
51. Mangione CM, Goldman L, Orav EJ, Marcantonio ER, Pedan A, Ludwig LE, et al. Health-related quality of life after elective surgery: Measurement of longitudinal changes. J Gen Intern Med 1997; 12: 686–697.
52. Nilsdotter AK, Aurell Y, Siosteen AK, Lohmander LS, Roos HP. Radiographic stage of osteoarthritis or sex of the patient does not predict one year outcome after total hip arthroplasty. Ann Rheum Dis 2001; 60: 228–232.
53. Nilsdotter AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41: 1261–1267.
54. O'Connell T, Browne C, Corcoran R, Howell F. Quality of life following total hip replacement. Ir Med J 2000; 93: 108–110.
55. O'Shea K, Bale E, Murray P. Cost analysis of primary total hip replacement. Ir Med J 2002; 95: 177–180.
56. Singer MA, Hopman WM, MacKenzie TA. Physical functioning and mental health in patients with chronic medical conditions. Qual Life Res 1999; 8: 687–691.
57. Croft P, Lewis M, Wynn JC, Coggon D, Cooper C. Health status in patients awaiting hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41: 1001–1007.
58. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78: 185–190.
59. Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of 'noise' in monitoring patient–based outcomes: Generic, disease-specific and site-specific instruments for total hip replacement. J Health Serv Res Policy 1996; 1: 224–231.
60. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. Med Care 2003; 41: 582–592.
61. Carmona L, Ballina J, Gabriel R, Laffon A. The burden of musculoskeletal diseases in the general population of Spain: Results from a national survey. Ann Rheum Dis 2001; 60: 1040–1045.
62. Cooper JK, Kohlmann T. Factors associated with health status of older Americans. Age Ageing 2001; 30: 495–501.
63. Wilcox S, Brenes GA, Levine D, Sevick MA, Shumaker SA, Craven T. Factors related to sleep disturbance in older adults experiencing knee pain or knee pain with radiographic evidence of knee osteoarthritis. J Am Geriatr Soc 2000; 48: 1241–1251.

64. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Health impact of pain in the hip region with and without radiographic evidence of osteoarthritis: A study of new attenders to primary care. The PCR Hip Study Group. Ann Rheum Dis 2000; 59: 857–863.
65. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: Comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci 2003; 8: 288–293.
66. Sprangers MA, de Regt EB, Andries F, van Agt HM, Bijl RV, de Boer JB, et al. Which chronic conditions are associated with better or poorer quality of life? J Clin Epidemiol 2000; 53: 895–907.
67. Wensing M, Vingerhoets E, Grol R. Functional status, health problems, age and comorbidity in primary care patients. Qual Life Res 2001; 10: 141–148.
68. Marks R, Allegrante JP. Comorbid disease profiles of adults with end-stage hip osteoarthritis. Med Sci Monit 2002; 8: CR305–CR309.

Address for correspondence: Caroline B. Terwee, Institute for Research in Extramural Medicine (EMGO Institute), VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands
Phone: +31-20-444-8187; Fax: +31-20-444-6775
E-mail: cb.terwee@vumc.nl