Rheumatology 2002;41:1155–1162

Influence of obesity on the development of osteoarthritis of the hip: a systematic review

A. M. Lievense, S. M. A. Bierma-Zeinstra, A. P. Verhagen, M. E. van Baar[1], J. A. N. Verhaar[2] and B. W. Koes

Departments of General Practice, [1]Public Health and [2]Orthopaedics, Erasmus Medical Centre, Rotterdam, The Netherlands

Abstract

Objective. To evaluate the evidence for the influence of obesity as a risk factor for the occurrence of osteoarthritis (OA) of the hip.

Methods. A bibliographical search of Medline, EMBASE and the Cochrane library until April 2000 was carried out. Articles describing studies of the relationship between obesity and the occurrence of hip OA were selected. The quality of the studies was assessed with a standardized set of criteria. The outcome of the studies was compared with respect to study characteristics and the quality score for the study. A best-evidence synthesis was used to summarize the results of the individual studies.

Results. Five longitudinal and seven cross-sectional studies were included in this review. There was no association between outcome and study design or methodological quality. The associations between obesity and hip OA were, however, stronger in studies in which the diagnosis of hip OA was based not only on radiological criteria but also on clinical symptoms. Overall, moderate evidence was found for a positive association between obesity and the occurrence of hip OA, with an odds ratio of approximately 2.

Conclusion. The evidence for a positive influence of obesity on the development of hip OA is moderate.

Osteoarthritis (OA) of the hip is a major cause of morbidity and disability in the elderly. This problem increases with the current ageing of the population in western societies. In addition to the pain and discomfort it causes, OA has major economic consequences [1].

Studies in Europe have estimated that approximately 7–25% of Caucasian individuals over the age of 55 yr suffer from hip OA; these estimates vary due to differences in the definition of OA or the selection of the study population. The prevalence of hip OA appears to be lowest in Asians, followed by African black and Native American populations, and is highest in white Europeans [2–5]. Over the last two decades many epidemiological studies have investigated the determinants of OA. These studies are important if we are to improve our understanding of the mechanisms leading to OA and to determine whether (modifiable) risk factors exist for which preventive interventions can be developed and investigated.

A frequently studied object of interest is weight, or body mass index (BMI), and its relationship with OA. A review published in 1998 showed that people with a higher BMI are more prone to the development of knee OA. However, due to sparse data and inconsistencies in the reported studies, the impact of obesity on hip OA was less clear [2].

Because several new studies investigating the relationship between hip OA and obesity have been published, we decided to investigate this topic using modern methods for systematically identifying and assessing the available studies. The result may be of considerable practical and theoretical importance for the management of this disease, including preventive measures.

Submitted 9 July 2001; revised version accepted 16 April 2002.

Correspondence to: A. Lievense, Department of General Practice, Erasmus University Rotterdam, Post-box 1738, 3000 DR Rotterdam, The Netherlands.

Methods

Identification and selection of the literature

To identify observational studies on this subject, relevant publications were searched using the following databases: Medline (1966 to April 2000), Cochrane library (1993 to April 2000) and EMBASE (1980 to April 2000). The following key words were used: [hip and (arthritis or arthrosis or osteoarthritis or osteoarthrosis) or coxarthrosis] and (risk factor or causative or determinants or predictor or aetiology) and (case–control or retrospective or prospective or longitudinal or follow

up or cohort). (A detailed list of the key words used can be obtained from the corresponding author.) We optimized the search strategy by looking at the specificity and sensitivity of different strategies. We tried to be as sensitive as possible within the bounds of feasibility, because of difficulty in finding the right key words and the different types of design. The search was extended by screening the reference lists of all relevant articles identified.

A study was eligible for inclusion if it fulfilled all of the following criteria: (i) one of the aims of the study was to investigate an association between hip OA and obesity; (ii) the article was written in English, Dutch, German, French, Danish, Norwegian or Swedish; (iii) the article was a full-text article; (iv) the patients in the studies had to suffer from radiological and/or clinical hip OA, had had a (total) hip replacement (THR) or were on the waiting list for one; and (v) the study was of a cohort, case–control or cross-sectional design.

A study was excluded if the studied population had a specific underlying pathology, such as trauma (fractures), infection, rheumatoid arthritis, ankylosing spondylitis, Perthes' disease, tuberculosis, haemochromatosis, sickle cell disease, Cushing's disease or femoral head necrosis.

Publication bias

Identification of all relevant articles is crucial to the validity of a systematic review [6]. The amount of potential publication bias in our study was analysed by means of a funnel plot, in which the study outcome was plotted against the sample size of the study. In the absence of publication bias, the plot will resemble a symmetrical inverted funnel [7]; because of the small number of studies included and the lack of the required data we could not perform a test for symmetry, so we visually examined the funnel plot for symmetry.

Methodological quality assessment

The methodological quality of observational studies can vary considerably, which may influence the results and conclusions of the studies at issue and consequently also the results and conclusions of a systematic review. Therefore, the quality of each included paper was assessed using the following method. Two reviewers (AML and SMABZ) independently scored the quality of the selected papers according to a standardized set of criteria (Table 1). These criteria have been used in previous reviews of observational studies in the field of musculoskeletal disorders [8–10] and were modified to cover the topic of our review. The criteria concern both the internal validity and the informativeness of the study. Only items reflecting the internal validity of the studies were used to assess the methodological quality.

In case of disagreement, both reviewers tried to achieve consensus; if the disagreement was not resolved, a third reviewer (BWK) was consulted in order to achieve a final judgement.

Several items are not applicable to certain types of study design (e.g. cohort or case–control study), and therefore do not contribute to the total score of that particular study. This means that the maximum score (100%) for each study was based only on the items applicable to that particular type of study design.

TABLE 1. List of criteria for the assessment of the methodological quality for cohort and case–control studies [8]

Item	Criterion	V/I[a]	CH/CC[b]
Study population			
1	Selection before disease was present or at uniform point	V	CH/CC
2	Cases and controls were drawn from the same population	V	CC
3	Participation rate $\geqslant$80% for cases/cohort	V	CH/CC
4	Participation rate $\geqslant$80% for controls	V	CC
5	Sufficient description of baseline characteristics	I	CH/CC
Assessment of risk factor			
6	Exposure assessment was blinded	V	CH/CC
7	Exposure was measured identical for cases and controls	V	CC
8	Exposure was assessed prior to the outcome	V	CH/CC
Assessment of hip OA			
9	Hip OA was assessed identical in studied population	V	CH/CC
10	Presence of hip OA was assessed reproducibly	V	CH/CC
11	Presence of hip OA was according to valid definitions [38, 39]	V	CH/CC
12	Classification was standardized [40–42]	I	CH/CC
Study design			
13	Prospective design was used	V	CH/CC
14	Follow-up time $\geqslant$3 yr	V	CH
15	Withdrawals $\leqslant$20%	V	CH
16	Information on completers *vs* withdrawals	I	CH
Analysis and data presentation			
17	Frequencies of most important outcomes were given	I	CH/CC
18	Appropriate analysis techniques were used	V	CH/CC
19	Adjusted for at least age and gender	V	CH/CC

[a]V is a criterion of validity/precision; I is a criterion of informativeness.
[b]CH, applicable to cohort studies; CC, applicable to case–control studies.

Positive scores were summed to give an overall internal validity score.

Best evidence synthesis

Because the observational studies were considered to be heterogeneous with regard to the population studied, methodological quality and determinants and outcome measures for hip OA, we followed standard practice and refrained from statistically pooling the data and performed a 'best evidence' synthesis [8, 11–14]. First, the studies were classified according to the type of study design. A prospective cohort study was judged as the preferred design, followed by a case–control study, and then by a cross-sectional study. After that, the studies were ranked according to their methodological quality score. The following ranking of the levels of evidence was formulated [8, 10, 11]:

1. Strong evidence is provided by generally consistent findings in multiple high-quality cohort studies.
2. Moderate evidence is provided by general consistent findings
   in one high-quality cohort study and two or more high quality case–control studies or
   in three or more high-quality case–control studies.
3. Limited evidence is provided by (general consistent) findings
   in a single cohort study,
   in one or two case–control studies or
   in multiple cross-sectional studies.
4. Conflicting evidence is provided by conflicting findings (i.e. $<75\%$ of the studies reported consistent findings).
5. No evidence is provided when no studies could be found.

A study was considered to be of high quality if the methodological quality score was $\geqslant 60\%$. This level was chosen at the mean of all quality scores [15].

Data extraction

Two researchers (AML and SMABZ) collected the characteristics of the included studies independently of each other. They collected items on the definition of the study population, how the presence or absence of hip OA was assessed, the assessment of obesity, if the study corrected for potential confounding factors and which results were reported.

When a study reported several outcomes because of a division of the study population into subgroups, the separate outcomes were combined (where possible) using Mantel Haenszel statistics, by methods described by Clayton and Hills [16] or the method described by Tan *et al.* [17] as appropriate.

Results

Identification and selection of the literature

A total of 2921 references were identified initially; of these, only nine articles met our selection criteria [4, 18–25]. The most frequent reasons for failing to meet our criteria were that there was not an appropriate study design (case reports, no data on the control group) or no specific information about the relationship between obesity and hip OA.

After screening the reference lists of the selected studies, another four studies were included [26–29]. All of them were indexed in Medline but used other descriptions of the study design. Hartz *et al.* used the term 'health survey' [26] and van Saase *et al.* used the term 'epidemiological survey' [28]. The other two studies did not specify the study design [27, 29]. We tried to optimize the search strategy by deleting the type of study design as one of the required key words. This improved the sensitivity, but the specificity would then have been too low (4431 instead of 1513 hits in Medline for four more articles). For one study, there were two publications reporting different aspects of the study [22, 24]. Both publications were used to extract data regarding the methods used and the results reported. Thus, finally, 12 studies were included in this review.

Publication bias

To investigate the amount of publication bias for our study, a funnel plot was made (Fig. 1). The plot shows the relationship between the distribution of the point estimates of the association between obesity and hip OA, and the sample size (n). Although the plot shows a more or less equal distribution, there is a lack of small studies showing an inverse association.

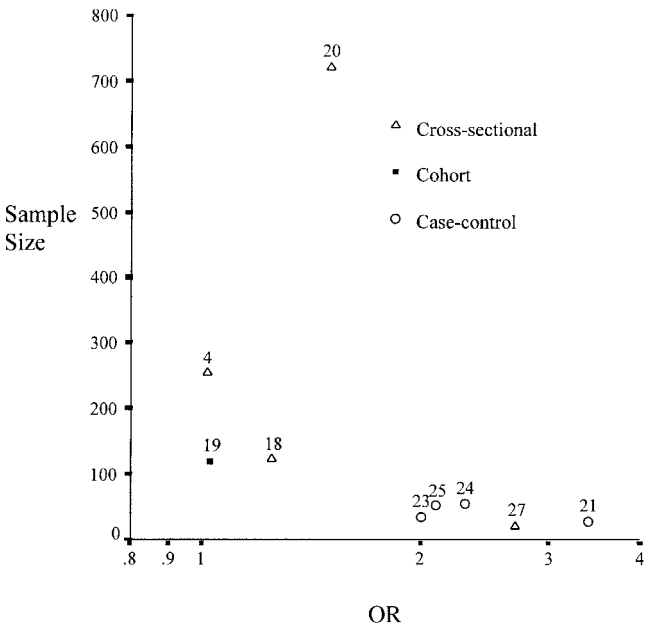

FIG. 1. Funnel plot showing the distribution of OR according to sample size. Of the 12 studies included, only nine provided sufficient data to be included in this plot. Three studies could not be included as no OR was available [26, 28, 29]. The numbers in the body of the graph are the reference numbers of the articles.

Table 2. Details of the studies included in this review

Author	Population	Assessment of hip OA	Adjusted for	Results[b]	Score
Cohort studies					
Gelber [19]	Male medical students (USA) classes 1948–1964 age 20–29 yr (n = 1180) Follow-up 36 (31–47) yr	Clinical (self-reported at follow-up by questionnaire)	Age, physical activity, joint injury	BMI <23, OR 1 (index) BMI 23–25, OR 1.10 (0.4–2.8) BMI >25, OR 1.03 (0.4–2.6) Data on age 40–49 yr	75
Case–control studies					
Vingard [24], Olsen [22]	1. All men from Stockholm with a THR between 1984 and 1988, aged 40–70 yr (n = 239) 2. Random selection of men from Stockholm aged 40–70 yr (n = 302)	[a]THR in that period	Age, smoking occupation, sports	BMI ⩽25, OR 1 (index) BMI >25, OR 2.3 (1.2–4.4) Data on age 50 yr	77
Oliveria [21]	1. Females of the Fallon Community Health Plan (USA) aged 20–89 yr with hip OA (n = 134) 2. Random selection of women from the Fallon Community Health Plan (n = 134)	[a]Clinical (ACR/pain, stiffness, soreness, aching, discomfort, swelling, tenderness) and X-ray (osteophytes, hypertrophic changes)	Age, gender, HRT, smoking, consumption of medical services	BMI ⩽24, OR 1 (index) BMI 24–28, OR 3.4 (0.4–25.6) BMI >28, OR 1.4 (0.1–17.5)	69
Vingard [25]	1. All women from South Sweden who had THR between 1992 and 1994 (n = 230) 2. Random selection of women from the same area without known hip disorders (n = 273)	[a]THR in that period	Age, number of children, sports, HRT, occupation, smoking	BMI ⩽20, OR 1 (index) BMI 20–25, OR 1.2 (0.5–2.8) BMI >25, OR 2.1 (0.9–4.6) Data on age 50 yr	69
Roach [23]	1. Males in Chicago with THR or signs of hip OA on X-ray (n = 99) 2. Same population without signs of hip OA on i.v. urograph (n = 233)	[a]Clinical (pain) X-ray/ i.v. urograph (JSW) or THR	Various confounders	BMI ⩽27.8, OR 1 BMI >27.7, OR 2.0 (1.0–4.0) Data on age 40 yr	46
Cross-sectional studies					
Cooper [18]	1. Patients on waiting list for THR at 2 clinics in England, aged >45 yr (n = 611) 2. Patients with the same GP, without previous hip surgery because of arthrosis (n = 611)	Waiting list for THR	Age, gender, Heberden's nodes, hip trauma	BMI ⩽25, OR 1 (index) BMI 25–28, OR 1.3 (0.95–1.7), bilateral OR 1.3 (0.9–1.8) BMI ⩾28, OR 1.6 (1.2–2.2), bilateral OR 1.7 (1.2–2.4)	69
Heliovaara [20]	Finnish open population (n = 7217), aged ⩾30 yr	Documented history of hip OA or definite physical findings	Age, gender, trauma, stress at work	BMI <25, OR 1 (index) BMI 25–30, OR 1.5 (1.1–1.9), bilateral OR 1.4 (1.0–2.0) BMI 30–35, OR 2.0 (1.5–2.7) bilateral OR 2.3 (1.5–3.5)	60
Van Saase [28]	Open population of Zoetermeer (The Netherlands), aged 45–64 yr (n = 2168)	X-ray (Kellgren and Lawrence)	Age, gender	'Association is absent'	60
Hartz [26]	Open population of the USA aged 40–69 yr (HANES I) (n = 4225)	X-ray (Kellgren and Lawrence)	Age, gender, race	'Relative weight was weakly associated with hip OA'	50
Tepper [4]	Open population of USA aged 55 yr (NHANES-1) (n = 2358)	X-ray (Kellgren and Lawrence)	Gender, race, age, education	BMI ⩽27, OR 1 (index) BMI >27, OR 1.02 (0.6–1.7), bilateral OR 2.0 (0.97–4.2)	50
Kraus [27]	1. Patients referred by physicians with hip OA to one clinic in USA (n = 100) 2. Patients at the same hospital at department of surgery or general medicine (n = 100)	[a]Clinical (patients) X-ray?	Age, gender, race	<20% above ideal weight, OR 1 <20% above ideal weight, ⩾20% OR 2.7 (1.4–5.4)	46
Saville [29]	1. Patients with primary hip OA at a department of special surgery (USA), aged 15–78 yr (n = 121) 2. US normals for height and weight 3. Patients from the same department without hip OA (n = 141)	[a]Clinical (patients) X-ray (JSN and subchondral bone sclerosis)		'Body weight distribution among patients with hip OA was similar to that of normal men and women in the US'	15

[a]Assessment of hip OA was only carried out for cases. [b]Figures in parenthesis are the confidence intervals. 1, cases; 2 and 3, control group.
NHANES, National Health and Nutrition Examination Survey; JSN, joint space narrowing; JSW, joint space width; i.v., intravenous; HRT, hormone replacement therapy.

Description of the studies

Table 2 gives a detailed description of the characteristics of the included studies. Only five studies had a longitudinal design, viz. one prospective cohort study [19] and four retrospective case–control studies [21–23, 25], whereas the remaining seven studies reported only cross-sectional associations between obesity and the presence of OA [4, 18, 20, 26–29]. In two of the 12 studies, the studied population was hospital-based [23, 27]. The other 10 studies were population-based.

Three studies included only males [19, 22, 23] and two others included only females [21, 25]. The ages of the studied populations diverge, but most studies investigated subjects aged 40 yr and older. All studies were carried out in the USA [4, 19, 21, 23, 26, 27, 29] or in Northern Europe [18, 20, 22, 25, 28]. Most of the studies determined obesity with the BMI. One study used the ideal body weight [27] and another study used relative weight [26]; both measurements were derived from a normal distribution of height and weight in the population [30]. One study used weight only [29].

The method of assessment of hip OA varied in several studies. Two studies relied on clinical information only [19, 20], three other studies used an X-ray score only [4, 26, 28] and another three studies used a combination of these two [21, 23, 29] as an outcome measure. Three studies used a (total) hip replacement (THR) or waiting for a THR in a specific period as an outcome [18, 22, 25]. One study did not describe clearly how hip OA was assessed; i.e. whether it was clinical only or also used X-ray information [27].

Results of the studies included

Nine of the 12 studies showed exact data of the outcomes; the other three described the outcome in global terms only [26, 28, 29].

Seven studies showed a positive association between obesity and hip OA [odds ratio (OR) $\geqslant 1.25$] [18, 20, 21, 23–25, 27]; in five studies the association was statistically significant [18, 20, 23, 24, 27], indicating that subjects with a higher BMI (approximately >25) have an increased risk of developing hip OA. Four of these seven studies had a case–control design and the other three a cross-sectional design. Three studies showed a weak positive relationship [relative risk (RR) or OR 1–1.25] [4, 19, 26]; one of these studies had a prospective cohort design.

There were no studies reporting a negative association. There were four studies reporting outcomes on the association between moderate obesity (BMI >25) and hip OA, but also between more severe obesity (BMI >27) and hip OA [18, 20, 21, 25]. Three of these studies showed a clear dose–response relationship [18, 20, 25]. Three studies also investigated the relationship of the BMI and the presence of bilateral disease [4, 18, 20]. The outcomes show that the strength of the relationship was similar to that found for unilateral hip OA.

In contrast with the nine studies which used clinical information (e.g. physical complaints, need for a THR) for the diagnosis of hip OA, the three studies using only an X-ray as evidence for hip OA [4, 26, 28] reported no association or only a very weak positive association with hip OA.

Methodological quality assessment

The two reviewers scored 404 items and agreed on 378 items (94%, $\kappa = 0.87$). The 26 disagreements were resolved in a single consensus meeting. Table 3 shows the studies in order of their methodological quality score, subdivided into the different types of study design (i.e. cohort, case–control and cross-sectional). The

TABLE 3. Quality scores of the studies

	Criterion																					
	1	2	3	4	5[a]	6	7	8	9	10	11	12[a]	13	14	15	16[a]	17[a]	18	19	Individual score	Total obtainable	Total score
Cohort studies																						
Gelber [19]	1	na	1	na	1	1	na	1	1	0	0	0	1	1	1	0	1	0	1	9	12	75%
Case–control studies																						
Olsen [22], Vingard [24]	1	1	1	1	1	0	1	1	0	1	1	1	0	na	na	na	1	1	1	10	13	77%
Oliveria [21]	1	1	1	1	0	0	1	1	0	0	1	1	0	na	na	na	1	1	1	9	13	69%
Vingard [25]	1	1	1	1	1	0	1	0	0	1	1	1	0	na	na	na	1	1	1	9	13	69%
Roach [23]	0	1	0	0	1	0	1	0	0	1	1	1	0	na	na	na	1	1	1	6	13	46%
Cross-sectional studies																						
Cooper [18]	1	1	1	0	1	0	1	0	1	1	1	1	0	na	na	na	1	1	1	9	13	69%
Heliovaara [20]	0	na	1	na	1	0	na	0	1	1	1	0	0	na	na	na	1	1	1	6	10	60%
van Saase [28]	0	na	0	na	0	1	na	0	1	1	1	1	0	na	na	na	1	1	1	6	10	60%
Hartz [26]	0	na	0	na	1	0	na	0	1	1	1	1	0	na	na	na	1	1	1	5	10	50%
Tepper [4]	0	na	0	na	1	0	na	0	1	1	1	1	0	na	na	na	0	1	1	5	10	50%
Kraus [27]	1	1	1	1	1	0	1	0	0	0	0	0	0	na	na	na	1	0	1	6	13	46%
Saville [29]	1	0	0	0	0	0	1	0	0	0	0	0	0	na	na	na	1	0	0	2	13	15%

Each item was scored 1 when it met the specified criteria listed in Table 1. If it did not meet the criteria or was not described at all, a score of 0 was assigned. Positive scores were summed to give an overall internal validity score.

[a]Informativity item; not included in the analysis.

na, not applicable.

scores ranged from 15 to 77% of the maximum obtainable score for each study design. The mean quality score was 57%.

The participation rates of the studied populations (Table 1, items 3, 4) were very low; the criteria were fulfilled by three of the four case–control studies and the cohort study but by only one out of the seven cross-sectional studies. In contrast with the assessment of obesity, for which there was almost always an identical or comparable assessment within the different groups of the population (Table 1, item 7), the assessment of hip OA (item 9) diverged substantially. In studies with a cohort as study base (longitudinal and cross-sectional data), there was always an identical assessment of hip OA. In case–control studies (with longitudinal and cross-sectional data), however, the assessment of hip OA in six of the seven studies was carried out in the patients but not (or differently) in the controls [21–23, 25, 27, 29].

One of the main concerns about the internal validity of observational studies is that of confounding [31]. When there is confounding, the observed effect may occur due to a combination of several determinants that are prevalent in the study population. In hip OA two potential confounders are age and gender, and there should be correction for at least these determinants in the individual studies. Only one study did not meet this criterion [29]; all the others did, by means of matching, restriction or adjustment in the analysis.

No correlation was found between the quality score and the study outcome (Fig. 2) (Pearson correlation $= -0.04$, $P = 0.9$).

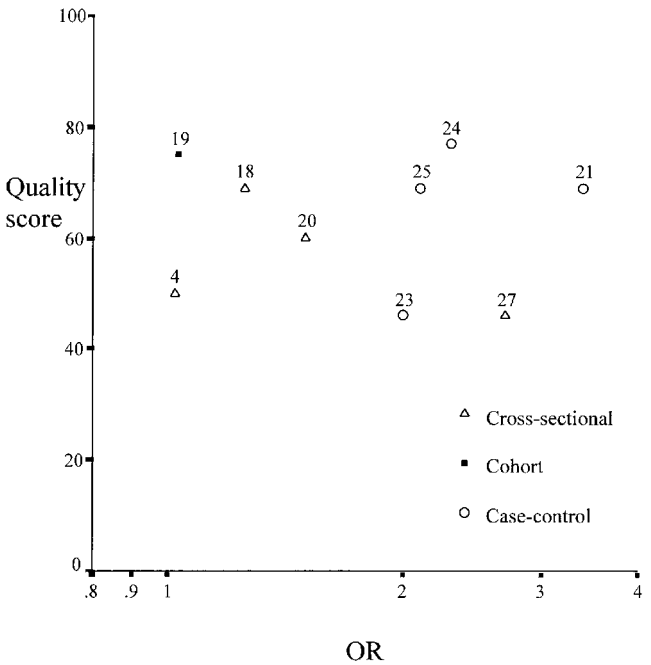

FIG. 2. Quality score of nine of the 12 studies in relation to OR. Three studies could not be included as no OR was available [26, 28, 29]. The numbers in the body of the graph are the reference numbers of the articles.

Three of the 12 studies could not be plotted due to a lack of precise data on the OR or RR. Of these, van Saase *et al.* [28] achieved a quality score of 67% and reported no association between obesity and hip OA. Hartz *et al.* [26] scored 56% and reported a weak association, whereas Saville and Dickson [29] achieved a score of 17% and showed no association between obesity and hip OA.

Best evidence synthesis

According to Chalmers and colleagues [15], the cut-off point for high-quality studies was set at the mean of the quality scores, so a study was considered to be of high quality if the methodological quality score was $\geqslant 60\%$.

There were five longitudinal studies, and four of them reached the level of high quality. The high-quality cohort study reported an OR of 1.03 [95% confidence interval (CI) 0.41–2.60], whereas the three high-quality case–control studies reported an association of 2.3 (95% CI 1.2–4.4), 3.4 (95% CI 0.4–25.6) and 2.1 (95% CI 0.9–4.6). This means that there is moderate evidence for a positive association between obesity and hip OA, with an OR of approximately 2.

In the subgroup of studies reporting the association between bilateral hip OA and obesity, only three cross-sectional studies were found, of which two reached the level of high quality [18, 20]. This means that, with the best-evidence synthesis, we can conclude that there is only limited evidence for a positive association between obesity and the occurrence of bilateral hip OA. The strength of the association is comparable to the strength found for unilateral hip OA.

For the subgroup of articles assessing hip OA on clinical information (e.g. physical complaints, need for a THR), six [18–22, 25] of the nine articles reached the level of high quality. One had a cohort design [19], three a case–control design [21, 22, 25] and two a cross-sectional design [18, 20]. The reported ORs were 1.03, 1.4, 2.3, 2.1, 1.6 and 2.0. This means that there is moderate evidence for a positive relationship between clinically assessed hip OA and obesity.

All three studies that assessed hip OA with radiological parameters had a cross-sectional design [4, 26, 28]. The reported outcomes were 'absence of an association', 'a weak association' and an OR of 1.02. One of them reached the level of high quality [32], so there is limited evidence for no association in the subgroup of radiologically assessed hip OA.

Discussion

In this systematic review, we summarized the evidence available in the literature on the influence of obesity on the development of OA of the hip. On the basis of this evidence, we may conclude that there is moderate evidence for a positive association between hip OA and obesity. For the subcategory of bilateral hip OA, only

limited evidence was found. The strength of association found in this subgroup was comparable to the strength found in the overall conclusion.

For a subcategory of articles in which hip OA was assessed clinically, moderate evidence was also found for a positive association, in contrast to radiologically assessed hip OA, for which limited evidence for no association was found. This review may, however, suffer from several restrictions.

Identification and selection of the literature

Although we put much effort into identifying the relevant articles, our literature search might have had some limitations. Some relevant articles may have been missed because they used other key words or had unclear abstracts, and not all published articles are indexed in databases. Besides, we excluded articles written in languages other than English, Dutch, German, French, Danish, Norwegian and Swedish.

The presumed publication bias found in our results (Fig. 1) can be explained by a selection of studies that were published for their 'exciting' positive results that satisfy current dogma. When assessing the funnel plot, we have to keep in mind that three of the 11 studies could not be included in the analysis due to insufficient data to deduce a point estimate [26, 28, 29].

Unfortunately, only one prospective study of the association between obesity and hip OA could be found. This design is known to be the most valid type of observational studies. Having more data from this type of design would have allowed us to provide a more valid and precise conclusion. The decision to include cross-sectional studies is based on the results of several studies reporting a similarity between current BMI and historical BMI in patients with hip OA [25, 33, 34]. Our data support these findings. We acknowledge, however, that the strength of the evidence of cross-sectional studies is much lower than the evidence provided by prospective cohort studies and, to a lesser extent, case–control studies.

Quality assessment and best-evidence synthesis

The quality assessment was challenging because no tested and validated criteria lists have been published for observational studies in the field of OA. Also, limited data were found on performing a best-evidence synthesis with observational studies. Therefore we have presented them in a reproducible manner, and the criteria we used were relatively strict.

A very notable item in the quality assessment was the difference in the assessment of hip OA between the patients and controls in the case–control studies (either longitudinal or cross-sectional). Most studies took as cases all patients with hip OA and as controls a sample of the underlying population, but without screening them for hip OA. This means that some controls might have had hip OA, so these studies would be biased towards a null result.

There have been suggestions that controls do not need to be evaluated for the occurrence of the disease, but rather must represent the base population, which will include such cases [35]. Considering the fact that there are still contradictory opinions on this subject, we prefer the classical point of view, namely that controls must be free of the disease under study.

The most interesting finding in our study is the difference found between clinically assessed and radiologically assessed hip OA. Whereas the former group of studies shows that obese people suffer more from hip OA, the latter shows that there is no difference in the presence of hip OA, between obese and non-obese people. This may suggest that obese people suffer more from the same radiological degree of hip OA than non-obese people do. Thus, obese patients may have more hip complaints at an equal radiological stage, due to the weight on the hip joint, and therefore qualify earlier for THR than non-obese patients. Unfortunately, there are only a few studies reporting on weight loss in relation to knee OA, and none in relation to hip OA, to support this explanation [36, 37].

In view of the moderate evidence for an association, knowledge of the origin of the relationship between hip OA and obesity is important. Future research, especially well-designed prospective cohort studies with adequate follow-up time, will not only strengthen the conclusion of this review but might also throw light on the cause of this relationship.

Conclusion

This systematic review shows that in the literature, there is moderate evidence for a positive influence of obesity on the development of hip OA.

Acknowledgement

This study was supported by a grant from the Dutch Arthritis Association.

References

1. Hadler NM. Osteoarthritis as a public health problem. Clin Rheum Dis 1985;11:175–85.
2. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41:1343–55.
3. Hoaglund FT, Oishi CS, Gialamas GG. Extreme variations in racial rates of total hip arthroplasty for primacy coxarthrosis: A population-based study in San Francisco. Ann Rheum Dis 1995;54:107–10.
4. Tepper S, Hochberg MC. Factors associated with hip osteoarthritis: data from the First National Health and Nutrition Examination Survey (NHANES-I). Am J Epidemiol 1993;137:1081–8.
5. Lawrence RC, Hochberg MC, Kelsey JL *et al.* Estimates of the prevalence of selected arthritic and musculoskeletal diseases in the United States. J Rheumatol 1989;16:427–41.
6. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994;1:447–58.

7. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Br Med J 1997;315:629–34.
8. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Influence of work on the development of osteoarthritis of the hip: a systematic review. J Rheumatol 2001;28:2520–8.
9. Borghouts JA, Koes BW, Bouter LM. The clinical course and prognostic factors of non-specific neck pain: a systematic review. Pain 1998;77:1–13.
10. Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM. Physical load during work and leisure time as risk factors for back pain. Scand J Work Environ Health 1999;25:387–403.
11. Bronfort G. Spinal manipulation: current state of research and its indications. Neurol Clin 1999;17:91–111.
12. Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. J Clin Epidemiol 1995; 48:9–18.
13. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. J Am Med Assoc 1995;274:1800–4.
14. Zeegers MP, Heisterkamp SH, Kostense PJ, van der Windt DA, Scholten RJ. De praktijk van systematische reviews. VII. Het combineren van de resultaten van observationele onderzoeken. Ned Tijdschr Geneeskd 2000;144:1393–7.
15. Chalmers TC, Smith H Jr, Blackburn B *et al.* A method for assessing the quality of a randomized control trial. Control Clin Trials 1981;2:31–49.
16. Clayton D, Hills M. Statistical models in epidemiology. Oxford: Oxford University Press, 1993.
17. Tan FES, Zeegers MPA, Dorant E, van Den Brandt PA. An alternative method of combining exposure specific odds ratios in epidemiologic studies. Research report. Maastricht: Department of Methodology and Statistics, University of Maastricht, 1999.
18. Cooper C, Inskip H, Croft P *et al.* Individual risk factors for hip osteoarthritis: obesity, hip injury, and physical activity. Am J Epidemiol 1998;147:516–22.
19. Gelber AC, Hochberg MC, Mead LA, Wang NY, Wigley FM, Klag MJ. Body mass index in young men and the risk of subsequent knee and hip osteoarthritis. Am J Med 1999;107:542–8.
20. Heliovaara M, Makela M, Impivaara O, Knekt P, Aromaa A, Sievers K. Association of overweight, trauma and workload with coxarthrosis. A health survey of 7,217 persons. Acta Orthop Scand 1993;64:513–8.
21. Oliveria SA, Felson DT, Cirillo PA, Reed JI, Walker AM. Body weight, body mass index, and incident symptomatic osteoarthritis of the hand, hip, and knee. Epidemiology 1999;10:161–6.
22. Olsen O, Vingard E, Koster M, Alfredsson L. Etiologic fractions for physical work load, sports and overweight in the occurrence of coxarthrosis. Scand J Work Environ Health 1994;20:184–8.
23. Roach KE, Persky V, Miles T, Budiman-Mak E. Biomechanical aspects of occupation and osteoarthritis of the hip: a case–control study. J Rheumatol 1994; 21:2334–40.
24. Vingard E. Overweight predisposes to coxarthrosis. Body-mass index studied in 239 males with hip arthroplasty. Acta Orthop Scand 1991;62:106–9.
25. Vingard E, Alfredsson L, Malchau H. Lifestyle factors and hip arthrosis. A case referent study of body mass index, smoking and hormone therapy in 503 Swedish women. Acta Orthop Scand 1997;68:216–20.
26. Hartz AJ, Fischer ME, Bril G *et al.* The association of obesity with joint pain and osteoarthritis in the HANES data. J Chronic Dis 1986;39:311–9.
27. Kraus JF, D'Ambrosia RD, Smith EG *et al.* An epidemiological study of severe osteoarthritis. Orthopedics 1978;1:37–42.
28. van Saase JL, Vandenbroucke JP, van Romunde LK, Valkenburg HA. Osteoarthritis and obesity in the general population. A relationship calling for an explanation. J Rheumatol 1988;15:1152–8.
29. Saville PD, Dickson J. Age and weight in osteoarthritis of the hip. Arthritis Rheum 1968;11:635–44.
30. Pai MP, Paloucek FP. The origin of the 'ideal' body weight equations. Ann Pharmacother 2000;34:1066–9.
31. Davies HT, Williams FL. Confounded by confounding: separating association from causation. Hosp Med 1999; 60:294–7.
32. van Saase JL, van Romunde LK, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis 1989;48:271–80.
33. Rissanen A, Heliovaara M, Aromaa A. Overweight and anthropometric changes in adulthood: a prospective study of 17,000 Finns. Int J Obes 1988;12:391–401.
34. Kuczmarski RJ. Prevalence of overweight and weight gain in the United States. Am J Clin Nutr 1992;55(2 Suppl.):495–502.
35. Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 1998.
36. Felson DT, Zhang Y, Anthony JM, Naimark A, Anderson JJ. Weight loss reduces the risk for symptomatic knee osteoarthritis in women. The Framingham Study. Ann Intern Med 1992;116:535–9.
37. Toda Y, Toda T, Takemura S, Wada T, Morimoto T, Ogawa R. Change in body fat, but not body weight or metabolic correlates of obesity, is related to symptomatic relief of obese patients with knee osteoarthritis after a weight control program. J Rheumatol 1998; 25:2181–6.
38. Gunther KP, Scharf HP, Puhl W *et al.* Reproducibility of radiographic evaluation in osteoarthritis of the hip. Reproduzierbarkeit Der Rontgenologischen Beurteilung Von Coxarthrosen. Z Orthop Ihre Grenzgeb 1997;135:3–8.
39. Vignon E, Conrozier T, Piperno M, Richard S, Carrillon Y, Fantino O. Radiographic assessment of hip and knee osteoarthritis. Recommendations: recommended guidelines. Osteoarthritis Cartilage 1999;7:434–6.
40. Altman RD. The classification of osteoarthritis. J Rheumatol 1995;43(Suppl.):42–3.
41. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132:514–22.
42. Murphy WA Jr, Altman RD. Updated osteoarthritis reference standard. J Rheumatol 1995;43(Suppl.):56–9.