Is there a causal link between knee loading and knee osteoarthritis progression? A systematic review and meta-analysis of cohort studies and randomised trials

Marius Henriksen,1 Mark W Creaby,2 Hans Lund,3 Carsten Juhl,3,4 Robin Christensen1,3

ABSTRACT

Objective: We performed a systematic review, meta-analysis and assessed the evidence supporting a causal link between knee joint loading during walking and structural knee osteoarthritis (OA) progression. Design: Systematic review, meta-analysis and application of Bradford Hill’s considerations on causation. Data sources: We searched MEDLINE, Scopus, AMED, CINAHL and SportsDiscus for prospective cohort studies and randomised controlled trials (RCTs) from 1950 through October 2013. Study eligibility criteria: We selected cohort studies and RCTs in which estimates of knee joint loading during walking were used to predict structural knee OA progression assessed by X-ray or MRI. Data analyses: Meta-analysis was performed to estimate the combined OR for structural disease progression with higher baseline loading. The likelihood of a causal link between knee joint loading and OA progression was assessed from cohort studies using the Bradford Hill guidelines to derive a 0–4 causation score based on four criteria and examined for confirmation in RCTs. Results: Of the 1078 potentially eligible articles, 5 prospective cohort studies were included. The studies included a total of 452 patients relating joint loading to disease progression over 12–72 months. There were very serious limitations associated with the methodological quality of the included studies. The combined OR for disease progression was 1.90 (95% CI 0.85 to 4.25; I²=77%) for each one-unit increment in baseline knee loading. The combined causation score was 0, indicating no causal association between knee joint loading and knee OA progression. No RCTs were found to confirm or refute the findings from the cohort studies. Conclusions: There is very limited and low-quality evidence to support for a causal link between knee joint loading during walking and structural progression of knee OA. Trial registration number: CRD42012003253

Strengths and limitations of this study

- We performed the first systematic evaluation of the evidence for a causal link between mechanical loading of the knee during walking and structural progression of knee osteoarthritis.
- This systematic review, meta-analysis, and causation analysis find very limited evidence of a causal link between knee joint loading during walking and structural progression of knee osteoarthritis.
- Few studies were included in the meta-analysis since only a small number of studies have been conducted. Further research is needed.

BACKGROUND

Osteoarthritis (OA) is the most common form of arthritis affecting a large section of the population, and is a major cause of illness and disability.1 2 The knee joint is most frequently affected,3 and due to the knee’s crucial role in independent ambulation, knee OA leads to considerable disability affecting an individual’s participation in society and independent living. Because no cure is available and many people are affected, healthcare costs associated with knee OA are enormous—even higher than for the more high-profile diseases such as diabetes, cancer and cardiovascular diseases.4 With its prevalence on the rise, knee OA poses a substantial socioeconomic and public health burden.

Walking is the most common form of human locomotion, and most people walk every day, resulting in millions of steps per year. Biomechanical loading on the knee joint during walking is estimated to exceed 2–3 times body weight; thus, knee joint loading is unavoidable component of an
independent lifestyle. One widely accepted theory is that higher knee joint loading during walking is causally linked with accelerated structural knee OA progression. In accordance with this theory, unloading is advocated in an attempt to slow or halt disease progression. Unspecific measures of high joint loading, such as excess body mass and body mass index (BMI), have been associated with OA development and obesity is believed to act mainly through high mechanical loads, although systemic effects have also been indicated. However, higher body mass does not necessarily lead to higher joint loading during walking, and more specific measures of joint loading are necessary to apply.

Objective estimates of joint loading during walking can be obtained by three-dimensional gait analysis. Typically, the knee adduction moment (KAM) or biomechanical modelling of compression forces is used. The KAM has been of particular interest because it reflects the medial to lateral joint load distribution, with good face validity with respect to the relative prevalence of medial, as opposed to lateral compartment, tibiofemoral knee OA. Despite its common usage, a recent systematic review did not find consistent evidence to the fact that the KAM differs between those with and without knee OA.

The relationship between knee joint loading during walking and knee OA has been a focus of knee OA-related biomechanical research for 10–20 years. Studies reveal that symptomatic treatment of knee OA results in increases in knee joint loads during walking, which are unwanted according to the hypothesised structural consequences of increased joint loading. This may generate confusion among researchers, clinicians and patients, who are interested in this information to aid them in prevention, palliation and treatment strategies. Establishing causal relationship among modifiable factors, such as knee joint loading, and disease-specific measures, such as structural disease progression, is imperative in the generation of effective strategies.

Seminal epidemiological work by Hill proposed a set of considerations (the Bradford Hill criteria) to systematically evaluate the existence of a causal link between an exposure and a health outcome. These criteria have been previously employed to demonstrate the causal link between smoking and lung cancer, and between dietary factors and coronary heart disease, and have the capacity to evaluate the link between knee joint loading and structural knee OA progression.

Therefore, the objectives of this study were (1) to systematically evaluate the evidence supporting a causal link between exposure to knee joint loading during walking and structural progression of knee OA based on Hill’s considerations on causality and (2) to determine which knee joint load variables had been studied sufficiently in randomised controlled trials (RCTs) and found to support the findings of prospective cohort studies.

**METHODS**

The methods of the study search strategy, inclusion criteria and data analysis were prespecified in a protocol (see online supplementary file 1) and pre-registered (PROSPERO 2012:CRD42012003253).

We searched MEDLINE, Scopus, AMED, CINAHL and SportsDiscus for prospective cohort studies and RCTs from 1950 through October 2013. Search strategies are available as an online supplementary file 2. The reference lists of the retrieved articles were also searched for additional cohort studies and RCTs. Two reviewers (MH and MWC) independently assessed study eligibility. Excluded studies and reasons for exclusion were recorded. Disagreement was resolved by discussion and consensus. We included original full-length articles pertaining to the association between knee joint loading during walking, measured by three-dimensional gait analysis, and structural disease progression assessed quantitatively or semiquantitatively by X-ray or MRI. We considered only those studies that included participants diagnosed with knee OA that were followed for at least 1 year (necessarily because of the sensitivity of imaging and slow disease progression). Cohort studies had to include estimates of knee joint loading during walking at baseline and imaging-based assessment of structural disease progression. Clinical trials had to be randomised and compare image-based structural disease progression among different knee joint loading interventions (eg, increase vs decrease), placebo or control.

**Data extraction and analysis**

The following data were extracted from the studies: (1) Study design; (2) Country of origin; (3) Number of subjects; (4) Characteristics of the subjects; (5) Type of knee joint loading estimate; (6) Magnitudes of the estimate; (7) Description of the interventions (if any) and (8) Structural disease progression outcome (ie, semiquantitative or quantitative measures on X-ray and/or MRI).

For prospective cohort studies, we extracted estimates of the association between baseline knee joint loading and structural disease progression from baseline to follow-up. A progression group is typically compared to a non-progression group at baseline and the results are typically reported as ORs for disease progression for a one-unit higher baseline knee joint loading. If no OR was reported, we either calculated it from the reported data (if possible), or contacted the study authors and requested for the OR. For the RCTs we aimed to compare relative risks of structural disease progression between knee joint loading modification and control groups.

The individual study results (ie, log OR values and their corresponding SEs) were combined in a random effects meta-analysis model using the generic inverse variance outcome type in Review Manager. Heterogeneity between trials was assessed using the standard Q-test statistic (testing the hypothesis of
homogeneity), and we present the I² value, which can be interpreted as the percentage of total variation across the studies due to heterogeneity.²⁶

**Methodological quality and risk of bias assessment**
The methodological quality of the cohort studies was assessed using a published check list.²⁷ Two of the authors (MH and MWC) assessed this individually and judged each criterion to be ‘Adequately described’, ‘Unclear’, or ‘Inadequately described’—corresponding to ‘low risk of bias’, ‘unclear risk of bias’, and ‘high risk of bias’, respectively. Disagreement was resolved by discussion. If an included study was authored by one or more of the current authors, a third reviewer (HL) was asked to perform a quality assessment. One quality assessment item was omitted (“Was a dose–response relationship between exposure and outcome demonstrated?”) because this item relates to the findings of the study and not the methodological quality, and assessment of dose–response forms part of the Bradford Hill criteria for causation (see below). The overall extent of risk of bias and methodological quality within each study was assessed using the GRADE approach to evaluate study limitations.²⁸

**Evaluation of evidence for causality**
Based on the Bradford Hill considerations on causality,²¹ a causation score was developed and used to systematically evaluate the evidence of a causal link between knee joint loading during walking and structural progression of knee OA. A similar score has previously been derived from the Bradford Hill considerations to assess the causal relationship between dietary factors and coronary heart disease.²³ The following criteria were used in the review of the cohort studies and given a score of 1 (criterion satisfied) or 0 (criterion not satisfied):

1. **Strength of association**: Associations quantified as a pooled OR ≥5.0, with lower 95% CI above 2.0; the expected direction was defined as ‘strong association.’ ‘Moderate association’ is defined as any statistically significant pooled OR (p<0.05). A statistically non-significant pooled OR was defined as ‘no association.’

2. **Consistency across studies**: An association requires replication in other studies. Consistency is defined as ≥75% of associations being strong or moderate.

3. **Temporality**: Refers to temporal relationship of association between exposure and outcome; exposure has to precede outcome. It is difficult to ensure temporal correctness because study participants are assumed to walk daily and are, therefore, exposed to knee joint loading throughout observation periods. We retained this criterion because temporality is necessary to infer causation; absence of temporal relationship between exposure and outcome precludes a causal link. Our analysis accepted studies with a temporarily correct design defined as baseline knee joint loading related to disease progression over time from that baseline. The temporality criterion was satisfied when temporal correctness was accepted in ≥75% of the included studies.

4. **Biological gradient**: When rate of progression increases (or decreases) incrementally as dose of exposure increases; provides strong evidence of causal relationship. This criterion is satisfied when ≥75% of tests for a trend pertaining to structural progression outcomes are statistically significant in the expected direction for knee joint loading.

The aforementioned four criteria scores were applied to derive one causation score for each knee joint loading variable. The scores were computed as the unweighted sum of the scores from each of the above criteria, for a possible range of 0–4. A score of 4 is considered strong evidence of a cause-and-effect relationship. A score of 3 is deemed to indicate moderate evidence. A score of 2 or less is considered weak evidence of causation.

A fifth criterion, **experimental evidence**, was used to examine whether the evidence from the prospective cohort studies was consistent with that from RCTs. Experimental evidence enhances the probability of causation, and may be used to upgrade or downgrade the calculated causation score.

The following four criteria were omitted:

- **Coherence**: Causation is more likely if what is observed is supported by and in agreement with the natural history of the disease. This criterion is usually applied when the outcome is assessed by surrogate outcomes. This criterion is omitted because it is satisfied by default since imaging modalities are surrogate outcomes.

- **Plausibility** relates to the assessment of whether the association is plausible or not. This criterion is omitted because of the highly subjective nature of this criterion.

- **Specificity** relates to the specific response to the exposure. This criterion is omitted because OA is a multisstructure disease; an association between knee joint loading and disease progression in multiple structures does not preclude a possible causal relationship.

- **Analogy** relates to the possibility that existing similar association can support causation (eg. Does the same association exist for hip or hand OA? If so, causation may be supported). This criterion is omitted because this review focuses specifically on knee OA.

**RESULTS**

**Study selection and characteristics**
The search yielded 1078 potentially eligible studies (figure 1; references available in online supplementary file 3). Of those, we included five prospective cohort studies published between 2002 and 2013 and involving a total of 452 patients with knee OA²⁹–³³ and 0 RCTs. The diagnostic criteria for knee OA used to determine participant inclusion in individual studies varied. The American College of Rheumatology (ACR) criteria³⁴ were used in one study,³² whereas a combination of
clinical and radiographic signs of knee OA was used as criteria in three studies, and one study did not report the diagnostic criteria used. Study characteristics are summarised in Table 1.

Joint loading and structural progression measures
Four cohort studies focused on the same knee joint loading estimate: the peak KAM. In these studies, peak KAM was assessed at baseline and related to structural medial disease progression assessed after a median of 15 months from that baseline (12–72 months). In one study, the baseline KAM impulse (area under the curve) was also related to medial disease progression. In the fifth study, overall knee compression loading was used as load exposure. Two studies defined structural disease progression as any worsening in semiquantitative radiographic medial joint space grade. One study defined structural disease progression as any loss of medial femoral cartilage volume (quantitative MRI) above a measurement error previously established. In another study, structural disease progression was assessed from semiquantitative grading of tibiofemoral cartilage loss and bone marrow lesions. In one study, structural progression was assessed semi-quantitatively as medial tibiofemoral cartilage loss and medial tibiofemoral bone marrow lesions, and quantitatively as medial tibial cartilage volume loss. For the meta-analysis, we extracted the association (OR) between peak KAM and semiquantitative progression in medial tibiofemoral cartilage loss. No ORs were available for the association between KAM and cartilage volume loss because both were continuous variables. However, the linear regression analysis was extracted for the assessment of the biological gradient criteria.

Association between joint loading and progression
ORs were not reported in three studies, but in individual patient data were reported and an OR was calculated. We contacted the authors of ref. asking them to provide an OR if possible. The authors returned a conference abstract including an OR published 4 years prior to the article. The abstract recorded 64 patients in contrast to the 57 patients included in ref. The authors could not account for the difference in number of the subjects between the full paper and conference abstract. One study grouped participants based on changes in joint loading following weight loss (increased vs reduced loads) and found no between group differences in structural disease progression. Owing to the use of an intervention (weight loss), it was not appropriate to extract OR for baseline loading.

The ORs for structural disease progression ranged from 0.40 to 6.46 for each unit higher baseline peak KAM magnitudes. From the random effects model, the combined OR was estimated to 1.90 (95% CI 0.85 to 4.25). The confidence in this estimate was downgraded due to heterogeneity (I²=77%). Figure 2 summarises the individual and pooled estimates.

In ref, the OR for progression of bone marrow lesions was 1.31 (95% CI 0.86 to 1.98) with every unit increment in baseline peak KAM; ORs for progression of cartilage defects and bone marrow lesions were 0.42 (95% CI 0.12 to 1.48) and 1.8 (95% CI 0.63 to 5.17), respectively, with every unit increment in baseline KAM impulse.

Methodological quality of the studies
One or more criteria were limited in the assessments of risk of selection, detection and attrition biases. Individual risk assessments are presented in online
supplementary appendix A. Across the studies, the comparability of the groups at baseline was generally inadequately or unclearly described (selection bias). Also, descriptions of how reliably the joint load exposures were ascertained were not included in four of the five studies (detection bias), and group differences in loss of subjects were not accounted for (attrition bias). The discrepancy in subject numbers in reports from one cohort,31 35 induced a high risk of selection and attrition bias. Table 2 presents summaries of the quality assessment. Agreement between quality raters was good ($\kappa$=0.73 (95% CI 0.48 to 0.97)).

Causality assessment

The non-significant combined OR indicates no association between knee joint load during walking and structural disease progression. Thus, a score of 0 in the Strength of Association criterion was assigned. The associations were not consistently replicated (50%), resulting in a score of 0 in the Consistency criterion. This finding was further supported by the heterogeneity between the studies ($I^2$=77%; figure 2), and the discrepancies between baseline peak KAM magnitudes among progressors and non-progressors across the cohorts (figure 3). Temporal correctness was accepted in one cohort.30 After requesting for additional information from the authors of one cohort study,31 we discovered that the article reported only part of the study. The additional information retrieved35 showed that the study consisted of three measurements taken every 18 months. Joint space and Kellgren-Lawrence grades (X-rays) were recorded at all three time points, whereas the KAM was obtained at time points 2 and 3 (18 and 36 months). KAM at time point 2 (18 months) was associated with structural disease progression from 18 to 36 months and structural disease progression from time points 1 to 2
(baseline to 18 months) was associated with increased KAM from 18 to 36 months.\textsuperscript{35} Thus, temporal correctness cannot be assigned to the observed associations in this study, and the overall Temporality criterion was not satisfied (a 0 score).

Biological Gradient was investigated in one cohort\textsuperscript{29} that showed no positive association between baseline peak KAM and loss of medial tibial cartilage volume over 12 months, resulting in a score of 0. Consequently, the causation score based on the included cohort studies was 0 for the association between peak KAM during walking and structural disease progression in patients with knee OA (table 3).

For the KAM impulse and overall knee compression loading, only one study was available for each exposure,\textsuperscript{29} \textsuperscript{33} and therefore no causation scores were calculated for these loading exposures.

As the systematic literature search did not identify any RCTs, the findings from the cohort studies can neither be confirmed nor refuted by experimental evidence. Hence, the fifth criterion could not be used to upgrade or downgrade the causation scores.

**DISCUSSION**

This review is the first to systematically assess—using systematic review, meta-analysis, and the Bradford Hill considerations on causality—whether a causal link exists between knee joint loading during walking and structural disease progression of knee OA. The systematic literature search identified five prospective cohort studies focusing on three estimates of knee joint loading: the peak KAM, KAM impulse and overall tibiofemoral compression force. Using a predefined algorithm, a causation score of 0 was reached, showing that there is no evidence of a causal link between knee joint loading during walking and structural progression of knee OA.

---

**Table 2** Summary of the quality assessment (risk of bias)

<table>
<thead>
<tr>
<th>Bias type</th>
<th>Quality criteria</th>
<th>Miyazaki \textit{et al}\textsuperscript{30}</th>
<th>Chang \textit{et al}\textsuperscript{\textsuperscript{1} \textsuperscript{35}}</th>
<th>Bennell \textit{et al}\textsuperscript{\textsuperscript{29}}</th>
<th>Woollard \textit{et al}\textsuperscript{\textsuperscript{22}}</th>
<th>Henriksen \textit{et al}\textsuperscript{\textsuperscript{33}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>Were the descriptions of the groups and the distribution of prognostic factors sufficient?</td>
<td>A</td>
<td>I</td>
<td>I</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Were the groups assembled at a similar point in their disease progression?</td>
<td>A</td>
<td>I</td>
<td>I</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Were the groups comparable on all important confounding factors?</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>U</td>
<td>A</td>
</tr>
<tr>
<td>Detection</td>
<td>Was the joint load estimate reliably ascertained?</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Was adequate adjustment made for the effects of these confounding variables?</td>
<td>A</td>
<td>I</td>
<td>A</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Was outcome assessment blind to exposure status?</td>
<td>A</td>
<td>A</td>
<td>I</td>
<td>I</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Was follow-up long enough for the outcomes to occur?</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>What proportion of the cohort was followed-up?\textsuperscript{*}</td>
<td>U</td>
<td>A</td>
<td>U</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Attrition</td>
<td>Were dropout rates and reasons for drop-out similar across groups?</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Overall risk of bias\textsuperscript{†}</td>
<td>Very serious limitations; high risk of bias</td>
<td>Very serious limitations; high risk of bias</td>
<td>Very serious limitations; high risk of bias</td>
<td>Very serious limitations; high risk of bias</td>
<td>Very serious limitations; high risk of bias</td>
<td></td>
</tr>
<tr>
<td>Methodological quality</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{*}Adequate=follow-up proportion >80%; unclear 50–80%; inadequate=<50%.

\textsuperscript{†}Risk of bias within studies is assessed using GRADE’s approach to study limitations: No serious limitation defined as all criteria being adequately described (high methodological quality); serious limitations defined as one criterion being inadequately described or >1 criterion being unclearly described (moderate methodological quality); very serious limitation defined as >1 criterion being inadequately described (low methodological quality).

A, adequately described; I, inadequately described; U, unclear.
Furthermore, no experimental data from RCTs exist to support or oppose the findings from the cohort studies. One of the prospective cohort studies compared increased versus decreased loads, without demonstrating differences in structural disease progression opposing a causal link.

Indirectly supporting the current findings, a longitudinal randomised trial comparing laterally wedged insoles (previously demonstrated to reduce loading) with placebo insoles, showed no group differences in structural disease progression. While that study seemingly provided experimental evidence, no joint load exposure data were included to document the assertion that joint loadings were in fact reduced in the ‘wedged insole’ group. Thus, the study did not fulfil the criteria for inclusion in this review. A recent meta-analysis on symptomatic effects of laterally wedged insoles showed no beneficial effects over neutral insoles, indicating limited effects of laterally wedged insoles. However, the null outcome is supported by the results in ref. 33, included in this review, showing neither detrimental effects of increased loading nor beneficial effects of decreased loading over 12 months in a group of knee OA patients.

The quality assessment of the included studies showed high risks of selection, detection and attrition biases (table 3). An important precondition for valid results in cohort studies is that groups should be comparable at baseline. Unless the groups are balanced for relevant baseline characteristics, differences in structural disease progression cannot confidently be attributed to the effects of interest (joint loads during walking). The included studies revealed significant inadequacies when assessing the comparability of the groups on relevant baseline characteristics, leading to prognostic imbalance. Four of the five cohort studies did not report how reliably the joint load estimates were ascertained (detection bias), which could influence the individual results. Moreover, the loss of subjects at follow-up in two of the five cohort studies was beyond 20%, with the dropout rates and reasons for dropout across groups not adequately accounted for (attrition bias). Finally, the discrepant numbers of subjects presented in different reports from the same cohort introduced a serious risk of bias and low confidence in the results. The overall low methodological quality of the included studies gives little confidence in the reported estimates of causation.

Despite the shortage of studies reporting a causal link between joint loading and structural disease progression, the conjectured relationship underlies the rationale for many clinical efforts to reduce knee loading during walking in attempts to retard disease progression. Weight loss, insoles, modified footwear, strengthening exercises, knee braces and gait modifications have been limited effects of laterally wedged insoles. However, the null outcome is supported by the results in ref. 33, included in this review, showing neither detrimental effects of increased loading nor beneficial effects of decreased loading over 12 months in a group of knee OA patients.

The quality assessment of the included studies showed high risks of selection, detection and attrition biases (table 3). An important precondition for valid results in cohort studies is that groups should be comparable at baseline. Unless the groups are balanced for relevant baseline characteristics, differences in structural disease progression cannot confidently be attributed to the effects of interest (joint loads during walking). The included studies revealed significant inadequacies when assessing the comparability of the groups on relevant baseline characteristics, leading to prognostic imbalance. Four of the five cohort studies did not report how reliably the joint load estimates were ascertained (detection bias), which could influence the individual results. Moreover, the loss of subjects at follow-up in two of the five cohort studies was beyond 20%, with the dropout rates and reasons for dropout across groups not adequately accounted for (attrition bias). Finally, the discrepant numbers of subjects presented in different reports from the same cohort introduced a serious risk of bias and low confidence in the results. The overall low methodological quality of the included studies gives little confidence in the reported estimates of causation.

Despite the shortage of studies reporting a causal link between joint loading and structural disease progression, the conjectured relationship underlies the rationale for many clinical efforts to reduce knee loading during walking in attempts to retard disease progression. Weight loss, insoles, modified footwear, strengthening exercises, knee braces and gait modifications have been limited effects of laterally wedged insoles. However, the null outcome is supported by the results in ref. 33, included in this review, showing neither detrimental effects of increased loading nor beneficial effects of decreased loading over 12 months in a group of knee OA patients.

The quality assessment of the included studies showed high risks of selection, detection and attrition biases (table 3). An important precondition for valid results in cohort studies is that groups should be comparable at baseline. Unless the groups are balanced for relevant baseline characteristics, differences in structural disease progression cannot confidently be attributed to the effects of interest (joint loads during walking). The included studies revealed significant inadequacies when assessing the comparability of the groups on relevant baseline characteristics, leading to prognostic imbalance. Four of the five cohort studies did not report how reliably the joint load estimates were ascertained (detection bias), which could influence the individual results. Moreover, the loss of subjects at follow-up in two of the five cohort studies was beyond 20%, with the dropout rates and reasons for dropout across groups not adequately accounted for (attrition bias). Finally, the discrepant numbers of subjects presented in different reports from the same cohort introduced a serious risk of bias and low confidence in the results. The overall low methodological quality of the included studies gives little confidence in the reported estimates of causation.

Despite the shortage of studies reporting a causal link between joint loading and structural disease progression, the conjectured relationship underlies the rationale for many clinical efforts to reduce knee loading during walking in attempts to retard disease progression. Weight loss, insoles, modified footwear, strengthening exercises, knee braces and gait modifications have been limited effects of laterally wedged insoles. However, the null outcome is supported by the results in ref. 33, included in this review, showing neither detrimental effects of increased loading nor beneficial effects of decreased loading over 12 months in a group of knee OA patients.

The quality assessment of the included studies showed high risks of selection, detection and attrition biases (table 3). An important precondition for valid results in cohort studies is that groups should be comparable at baseline. Unless the groups are balanced for relevant baseline characteristics, differences in structural disease progression cannot confidently be attributed to the effects of interest (joint loads during walking). The included studies revealed significant inadequacies when assessing the comparability of the groups on relevant baseline characteristics, leading to prognostic imbalance. Four of the five cohort studies did not report how reliably the joint load estimates were ascertained (detection bias), which could influence the individual results. Moreover, the loss of subjects at follow-up in two of the five cohort studies was beyond 20%, with the dropout rates and reasons for dropout across groups not adequately accounted for (attrition bias). Finally, the discrepant numbers of subjects presented in different reports from the same cohort introduced a serious risk of bias and low confidence in the results. The overall low methodological quality of the included studies gives little confidence in the reported estimates of causation.
suggested as means to lower knee joint loads and thus slow structural disease progression through mechanical pathways, with effects yet to be proven. The rationale for such efforts is diluted by the current results, not only by the low causality score, but also by the inconsistency across studies in the loading magnitudes between individuals who progress and those that do not. Figure 3 illustrates that patients with a certain baseline peak KAM magnitude in one study progress, whereas those with similar baseline peak KAM magnitudes in other studies do not. These differences in knee joint loading magnitudes between studies may be due to differences in gait analysis protocols across the involved gait labs, for example, in the selection of reference frame convention for the calculation of the KAM and footwear condition. However, including diversity of measures of the same variable can strengthen the confidence in the results of a systematic review.

The results from the included studies exhibited a high degree of heterogeneity (I² from the meta-analysis) (figure 2). The two cohorts with significant OR both used radiography as the structural outcome assessment, whereas the studies reporting non-significant ORs used MRI. Furthermore, the two studies reporting significant associations had longer follow-up periods (72 and 18 months), whereas the studies without significant associations assessed disease progression after 12 months. This concurrence of follow-up time and imaging modality precludes conclusions about whether the differences in results are based on follow-up time or imaging modality. Radiography is the current clinical gold standard for assessing structural disease progression, but it has low temporal sensitivity, making long follow-up times necessary. MRI—with its shorter temporal sensitivity—is possibly the more promising tool for assessing single structures and should be used in studies with longer follow-up times. The total number of patients was 452 in five cohorts; that is a very small number of patients when considering the large population and when compared to similar reviews in other chronic diseases. For instance, a recent review of the causal relationship between dietary factors and coronary heart disease included several millions of individuals in 361 cohorts and a multitude of dietary exposures.

Our study has strengths because we undertook several measures to minimise bias, restricting our review to studies with the strongest causal inference (cohort studies and RCTs), conducted independent assessments of study eligibility, and using predefined criteria to evaluate causation and methodological quality. We may be criticised for using arbitrary OR cut-offs to define strong and moderate association, but identifying the true cut-off for defining clinically meaningful OR values is impossible in a field that has few longitudinal studies to consult. An important limitation is the causality score that applied. The score is a modification of a previous similar score, thus it has limited empirical evidence to support its validity. Also, the causality score omitted several of the original causation considerations as proposed by Hill, yet the applied method was prespecified. We may also be criticised for analysing any image-based structural deterioration, rather than separating the analyses in X-ray and MRI-based assessments. However, the analysis plans were also prespecified, without prior knowledge of the number of studies and specifics of the outcomes.

Current guidelines on interpretation of data from observational studies and data combined from repeated observational studies (with a reasonable internal validity) would correspond to confidence in the estimate’s being ‘Low-quality evidence’, but could be upgraded if there is great confidence in the estimates and the effects are large. However, the current analysis revealed a large inconsistency among study results, poor precision of the combined analysis, and low methodological quality. Thus, we downgraded further to ‘very low quality evidence’. The inconsistency of the available studies may limit the appropriateness of meta-analyses, yet support the overall conclusion of this systematic literature review.

In conclusion, our systematic review, meta-analyses, and application of Bradford Hill’s considerations on causation shows that there is no evidence of a causal link between estimates of knee joint loading during walking and structural progression of knee OA. Future large well-designed prospective cohorts with subsequent confirmation from randomised trials are strongly recommended. The current study findings question the rationale behind clinical efforts to reduce knee joint loading during walking as such efforts are currently supported by very low-quality evidence showing no causal link between knee joint loading during walking and progression of knee OA.
non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES

Is there a causal link between knee loading and knee osteoarthritis progression? A systematic review and meta-analysis of cohort studies and randomised trials

Marius Henriksen, Mark W Creaby, Hans Lund, Carsten Juhl and Robin Christensen

*BMJ Open* 2014 4:
doi: 10.1136/bmjopen-2014-005368

Updated information and services can be found at:
http://bmjopen.bmj.com/content/4/7/e005368

These include:

**Supplementary Material**
Supplementary material can be found at:
http://bmjopen.bmj.com/content/suppl/2014/07/15/bmjopen-2014-005368.DC1

**References**
This article cites 46 articles, 11 of which you can access for free at:
http://bmjopen.bmj.com/content/4/7/e005368#B1BL

**Open Access**
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Epidemiology (2016)
- Evidence based practice (697)
- Rheumatology (155)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/