Dietary supplements for treating osteoarthritis: a systematic review and meta-analysis

Xiaoqian Liu,1,2 Gustavo C Machado,3 Jillian P Eyles,1,2,4 Varshini Ravi,1,2 David J Hunter1,2

ABSTRACT

Objective To investigate the efficacy and safety of dietary supplements for patients with osteoarthritis.

Design An intervention systematic review with random effects meta-analysis and meta-regression.

Data sources MEDLINE, EMBASE, Cochrane Register of Controlled Trials, Allied and Complementary Medicine and Cumulative Index to Nursing and Allied Health Literature were searched from inception to April 2017.

Study eligibility criteria Randomised controlled trials comparing oral supplements with placebo for hand, hip or knee osteoarthritis.

Results Of 20 supplements investigated in 69 eligible studies, 7 (collagen hydrolysate, passion fruit peel extract, Curcuma longa extract, Boswellia serrata extract, curcumin, pycnogenol and L-carnitine) demonstrated large (effect size >0.80) and clinically important effects for pain reduction at short term. Another six (undenatured type II collagen, avocado soybean unsaponifiables, methylsulfonylmethane, diacerein, glucosamine and chondroitin) revealed statistically significant improvements on pain, but were of unclear clinical importance. Only green-lipped mussel extract and undenatured type II collagen had clinically important effects on pain at medium term. No supplements were identified with clinically important effects on pain reduction at long term. Similar results were found for physical function. Chondroitin demonstrated statistically significant, but not clinically important structural improvement (effect size −0.30, −0.42 to −0.17). There were no differences between supplements and placebo for safety outcomes, except for diacerein. The Grading of Recommendations Assessment, Development and Evaluation suggested a wide range of quality evidence from very low to high.

Conclusions The overall analysis including all trials showed that supplements provided moderate and clinically meaningful treatment effects on pain and function in patients with hand, hip or knee osteoarthritis at short term, although the quality of evidence was very low. Some supplements with a limited number of studies and participants suggested large treatment effects, while widely used supplements such as glucosamine and chondroitin were either ineffective or showed small and arguably clinically unimportant treatment effects. Supplements had no clinically important effects on pain and function at medium-term and long-term follow-ups.

INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis, affecting about 12% of the population.1 The prevalence of OA is rising due to the ageing population and growing rates of obesity.2 Hip and knee OA was recently ranked as the 11th highest contributor to global disability,3 with enormous economic burden. Costs are estimated between 0.25% and 0.50% of a country’s gross domestic product, and the mean annual incremental direct and indirect costs are nearly US$7000 per person with OA.4

Paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs) are currently endorsed by clinical guidelines to treat OA,5–7 but emerging evidence has challenged this recommendation and revealed the potential for adverse events.8 9 This new evidence could potentially increase public interest in the benefits of alternative treatments. Currently 69% of patients with OA take some form of dietary supplements for their condition.10 They are regularly used by a wide range of consumers with musculoskeletal ailments and often used alongside or as an alternative to first-line interventions such as exercise. This has driven expenditures for alternative therapies to nearly equal those spent on traditional pharmacological therapy.11 Supplements account for US$25 billion sales annually,12 glucosamine and chondroitin together account for a third of this cost,13 with annual sales amounting to US$872 million.14 Although widely used by patients with OA, the safety and efficacy of supplements remain unclear and are often clouded by misinformation in mainstream media.15 16 Thus, physicians’ scepticism towards their efficacy can be challenged on this basis. Given the efforts of many to demonstrate the effect size (ES) of other first-line treatments such as exercise, it is critical that we have a clear understanding of the relative merits of supplement use. The aim of this review was to evaluate efficacy and safety of supplements compared with placebo in people with hand, hip or knee OA.

METHODS

The review, in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement,17 followed the Cochrane Handbook18 and the updated guideline of the Cochrane Musculoskeletal Review Group.19 The study protocol was registered on PROSPERO (https://www.crd.york.ac.uk/prospero/) (CRD42015025732).

Study selection

We included peer-reviewed, published, randomised placebo-controlled trials (RCTs) with full-text articles investigating the efficacy or safety of any dietary supplements for participants with hand, hip or knee OA. We defined dietary supplements as any supplement administered orally, and marketed as being beneficial for OA.
as a product taken orally intended to supplement the diet, which may include vitamins, minerals, herbs or other botanicals, amino acids, and substances. To be eligible for inclusion, trials were required to explicitly report including participants diagnosed with hand, hip or knee OA according to one of the following criteria: (1) American College of Rheumatology, (2) European League Against Rheumatism or (3) clinically and/or radiologically confirmed. Trials were required to report at least one of the following primary outcomes: pain or physical function. Secondary outcomes included adverse effects (AEs), radio logically confirmed. Trials were required to report at least one of the following primary outcomes: pain or physical function. Secondary outcomes included adverse effects (AEs), radio graphically joint space width/narrowing, stiffness, use of analgesics and quality of life. We were interested in evaluating individual supplement, so combination therapy and herbs were excluded. The language was limited to English.

Data sources and searches
We carried out a systematic electronic search on MEDLINE, EMBASE, Cochrane Register of Controlled Trials, Allied and Complementary Medicine and Cumulative Index to Nursing and Allied Health Literature from inception to 18 April 2017. Relevant keywords were used to construct the search strategy (with guidance from a librarian), including dietary supplements, nutraceuticals, OA and specific supplements (eg, glucosamine, chondroitin), according to previous reviews in which the supplements were reported for treatment of OA21–24 (online supplementary file 1, Search strategy). The reference lists of included studies were screened for potentially eligible trials. One reviewer (XL) conducted the first screening for potentially relevant records based on titles and abstracts. Full-text articles were independently assessed by two reviewers (XL and VR); disagreements were resolved through consensus or by consulting a third reviewer (DJH).

Data extraction and quality assessment
Baseline characteristics and outcomes data were extracted on a standardised form by two independent assessors. Factorial trials were assumed as parallel-group design.25 26 Data of multiple groups from one study were extracted using recommendations of the Cochrane Handbook.18 For continuous outcomes the following were extracted: means, SD and sample sizes at baseline and follow-up. If these were unavailable, change scores or mean differences were extracted. For dichotomous outcomes, the number of cases and total sample size were extracted. Safety outcomes included the number of participants reporting any or serious AEs, or withdrawn from the study because of AEs. We contacted authors for missing data. When authors were unav ailable, data were estimated using the recommendations of the Cochrane Handbook (eg, estimation of SD from SEs).18

We predefined a follow-up period ≤3 months as short term, a follow-up between 4 and 6 months as medium term, and a follow-up period >6 months as long-term. Data at the latest time point were used for those studies reporting multiple follow-up times within an individual time point category. When studies reported more than one scale to measure pain or physical function, we used the list of outcomes with hierarchy levels proposed by the Cochrane Musculoskeletal Review Group to choose the most relevant outcome.27

Risk of bias of included studies was assessed using the Cochrane Collaboration’s tool.18 If at least one of the bias domains was judged at high risk, the trial was considered at high risk; if all bias domains were judged at low risk, the trial was considered at low risk; otherwise, the trial was classified unclear risk. The role of the pharmaceutical companies in conducting the trial was assessed in the ‘other’ bias domain. We used RevMan V.5.3 to generate figures for the risk of bias assessment. Data extraction and assessment of risk of bias were performed independently by two reviewers (XL and VR). Disagreements were resolved through consensus or by a third reviewer (DJH).

Data synthesis and analysis
Data were pooled using random-effects model. As the included trials used different scales to assess the same outcome, standardised mean differences (SMD) were calculated for continuous data, and risk ratios (RR) were used for dichotomous data with their 95% CIs. To facilitate interpretation of estimated treatment effects, we considered ES up to 0.3 as small, between 0.3 and 0.8 as moderate, and greater than 0.8 as large effects.28 A threshold of 0.37 standardised units was defined as the minimum clinically important difference (MCID), which was based on the median MCID found in recent studies in patients with OA.9 29 We considered the effects as clinically important if ES and 95% CI exceeded our prespecified MCID threshold of 0.37. The effects were considered unclear if clinically important when the 95% CI crossed this threshold, and statistically significant but clearly not clinically important when the 95% CI was below the MCID but above the null effect line. If the 95% CI crossed the null effect line, supplements were considered ineffective.

Between-trial heterogeneity was assessed using the I² statistic; values higher than 50% were considered to indicate substantial heterogeneity.30 In order to explore potential sources of heterogeneity and the influence of different factors on the estimate of treatment effect, a series of sensitivity analyses were conducted for glucosamine and chondroitin. Meta-regression was used to compare the difference in ES (with 95% CI) and generate a p value between subgroups for pain at short term. Subgroups were defined in terms of judgements for each risk of bias domain (low, unclear or high); sample size of trials (large n≥50; small n≤50 per group); and different OA sites (eg, knee, hip). Statistical analyses were performed using Comprehensive Meta-Analysis V.3 (Englewood, New Jersey, USA).

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach was used to rate the quality of evidence for each pooled analysis according to five criteria: limitation of study design, publication bias, imprecision, inconsistency and indirectness.18 31 The quality of evidence was down-graded by one level for limitation of study design when more than a quarter of the studies included in an analysis were considered at high risk of bias. Where possible, publication bias was assessed by visual inspection of funnel plots (scatterplot of the ES from individual studies against its SE) and the Egger’s test for the meta-analysis with 10 or more trials.32 The rating was down-graded by one level if two-tailed p<0.1. Results were considered inconsistent if the heterogeneity between trials was substantial (I²>50%),33 and imprecise if the pooled sample size was <300 for dichotomous or <400 for continuous outcomes.34 The indirectness criterion was not considered as this review concerned a specific population, relevant outcomes and direct comparisons. GRADE ratings were defined as very low, low, moderate or high quality evidence reflecting the extent to which we are confident that an effect estimate is correct.

RESULTS
Initial search and result
We initially identified 2964 records, and 69 randomised placebo-controlled clinical trials with 11 586 participants investigating 20 supplements were included in this review...
Glucosamine (n=18 trials) and chondroitin (n=14 trials) were the most frequently investigated supplements, whereas eight supplements were investigated by only one trial. Fifty-eight of the 69 included trials (84%) investigated knee, five (7%) studied hip and knee, four (6%) assessed hip and two (3%) included hand OA. Sixty-seven of the 69 included trials (96%) were parallel-group along with two factorial trials (online supplementary file 1, Baseline characteristics). Sixty-six trials were included in the meta-analyses. Among three trials excluded from the analysis, one reported mean outcome values without dispersion measures; imputation of missing values was impossible, another reported data using percentages and lacked baseline scores for the placebo group, and another reported long-term outcomes as percentages.

Across trials, the mean age of participants ranged from 48 to 69 years, and the mean proportion of women was 65%. The treatment and follow-up duration ranged from 2 weeks to 3 years. Forty-four of the 69 included trials reported short-term outcomes (64%), 17 trials (25%) reported medium-term outcomes, and 17 trials (25%) reported long-term outcomes. A variety of proprietary products of supplements and different dosages were evaluated. Where supplements were judged similar (eg, curcumin and curcuminoid) or the same supplement was investigated using different preparations (eg, glucosamine sulfate and glucosamine...
Review

hydrochloride) or different dosages, we combined data according to recommendations from the *Cochrane Handbook*.

Overall, only 7 of the 69 included trials (10%) were judged at low risk of bias, 32 (46%) at high and 30 (44%) at unclear risk. The major potential source of bias was in the ‘other’ bias domain. Forty-four of the 69 included trials (64%) were funded by pharmaceutical companies. Twenty-six of the 44 pharmaceutically funded trials (59%) were judged at high risk of bias due to conflicts of interest (eg, employee/s was/were authors, or companies played an important role in conducting the trial). Incomplete outcome data and selective outcome reporting were other potential sources of bias in 17% and 6% of the 69 included trials, respectively. Online supplementary file 1 provides a graphical summary of the results of risk of bias for the 69 studies.

Primary analyses

Our pooled analysis suggested a moderate effect of supplements in reducing pain at short term (SMD $-0.53$, 95% CI $-0.61$ to $-0.46$, I² 78%) with very low-quality evidence. Seven supplements (L-carnitine, pycnogenol, curcumin, *Boswellia serrata* extract, *Curcuma longa* extract, passion fruit peel extract and collagen hydrolysate) showed large and clinically important effects for pain reductions at short term (figure 2). However, most of the studies were sponsored by pharmaceutical companies.

![Figure 2](http://bjsm.bmj.com/)

**Figure 2** Improvements in pain. Stratified according to supplement and time point. The order of supplements in the forest plots is to follow the effect size (ES) from small to large. The square data markers indicate the subtotal weighted standardised mean difference (SMD). The horizontal lines indicate 95% CIs. The diamond data markers represent overall weighted SMD and 95% CI. The solid vertical line shows the line of no effect (SMD=0). The dashed line shows the minimum clinically important difference (ES=$-0.37$).
companies and the research evidence was limited in the number of studies and participants included. Only one study was included for L-carnitine, *Curcuma longa* extract, passion fruit peel extract and collagen hydrolysate. Another six (avocado soybean unsaponifiables, undenatured type II collagen, methylsulfonylmethane (MSM), diacerein, glucosamine and chondroitin) revealed statistically significant improvements on pain at short term, but were of unclear clinical importance. The remaining five (vitamin E, bromelain, willow bark extract, *Artemisia annua* extract, green-lipped mussel extract) were ineffective in reducing pain compared with placebo at short term. However, green-lipped mussel extract and undenatured type II collagen demonstrated clinically important effects for pain reduction among eight supplements investigated at medium term. At long-term follow-up, none of the included supplements showed clinically important effects compared with placebo. The GRADE ratings ranged from very low-quality to high-quality evidence (online supplementary file 1).

Overall, there was very low-quality evidence of moderate physical function improvements provided by supplements at short term (SMD $-0.53$, 95% CI $-0.62$ to $-0.45$, $I^2$ 80%). We found that seven supplements (pycnogenol, passion fruit peel extract, *Curcuma longa* extract, L-carnitine, *Boswellia serrata* extract, curcumin and MSM) had large and clinically important effects on physical function at short term (figure 3). However, only one study was included for passion fruit peel extract, *Curcuma longa* extract, L-carnitine and curcumin. The research evidence is limited. Five supplements (undenatured type II collagen, avocado/soybean unsaponifiables (ASU), glucosamine, chondroitin and diacerein) demonstrated effects that were of unclear clinical importance, although these were statistically significant. Another four (vitamin E, bromelain, willow bark...

extract, *Artemisia annua* extract) were ineffective in improving physical function for OA at short term. At medium and long term, none of the included supplements revealed clinically important effects, with most being ineffective compared with placebo. The GRADE evaluation ranged from very low-quality to high-quality evidence (online supplementary file 1).

Pooled results suggested that five supplements (*Boswellia serrata* extract, L-carnitine, MSM,undenatured type II collagen and pycnogenol) revealed statistically significant improvements on stiffness at short term; however, only *Boswellia serrata* extract reached clinical importance (SMD −0.94, 95% CI −1.26 to −0.62, I² 0). No evidence existed to support any supplements improved stiffness at medium and long term. The GRADE ratings ranged from very low-quality to high-quality evidence (online supplementary file 1).

For the supplements (glucosamine, chondroitin, ASU, diacerein and vitamin D) included in studies of long-term structural improvement, only chondroitin demonstrated statistically significant difference (SMD −0.30, 95% CI −0.42 to −0.17, I² 5) when compared with placebo, but it did not reach clinically important effects. GRADE ratings were moderate to high (online supplementary file 1).

Limited evidence suggested that glucosamine, ASU and pycnogenol probably reduced the use of NSAIDs. Only *Curcuma longa* extract demonstrated statistically significant differences in analgesic use (RR 0.5, 95% CI 0.3 to 0.8). No significant differences in quality of life were demonstrated in five investigated supplements (glucosamine, chondroitin, MSM, vitamin E and willow bark extract) (online supplementary file 1).

For the safety analyses, we included 34 trials with 5271 participants including 12 supplements with regard to any adverse effect (figure 4). No statistically significant difference between supplements and placebo was found (RR 1.0, 95% CI 1.0 to 1.1), except for diacerein (RR 1.3, 95% CI 1.1 to 1.6) deriving from five trials with 1300 participants. Similarly, 34 trials with 8111 participants including 8 supplements revealed no difference in the number of dropouts due to adverse events (RR 1.2, 95% CI 1.0 to 1.4) and again except for diacerein (RR 1.7, 95% CI 1.2 to 2.3) deriving from 5 trials with 1360 participants. Only 9 trials with 2494 participants including 6 supplements reported serious adverse events data and there was no difference between groups (RR 1.3, 95% CI 1.0 to 1.6). The GRADE ratings ranged from low-quality to high-quality evidence (online supplementary file 1).

Inspection of the funnel plots (online supplementary file 2) suggested small study effects/publication bias for overall results in terms of pain and physical function at short and medium term. Similar results were found for glucosamine in terms of pain and physical function at short term. Egger’s test reached statistical significance (p<0.10).

**Sensitivity analyses**

For glucosamine (n=10 trials, table 1), trials judged at unclear risk of incomplete outcome (ES difference −1.0, 95% CI −1.5 to −0.4; p=0.001) or high risk of selective reporting bias (ES difference −0.9, 95% CI −1.5 to −0.4; p=0.001) revealed significantly larger treatment effects compared with those at low or unclear risk. Moreover, trials judged at unclear risk of allocation concealment suggested larger treatment effects compared with those at low risk; however, the difference was not statistically significant (ES difference −0.4, 95% CI −0.9 to 0.0; p=0.06). Industry-sponsored trials with conflicts of interest (high risk of ‘other’ bias) demonstrated larger effects than unfunded trials or trials without conflicts of interest (ES difference −1.4, 95% CI −2.4 to −0.4; p=0.003). Smaller trials reported larger beneficial effects compared with large trials (ES difference −0.5, 95% CI −0.9 to −0.2; p=0.003).

For chondroitin (n=9, table 2), trials with low risk of selection bias revealed smaller treatment effects than those judged at unclear risk (ES difference −0.4, 95% CI −0.6 to −0.1; p=0.005). Trials judged at unclear risk of allocation concealment or incomplete outcome suggested larger treatment effects compared with those at low risk, but...
the difference was not statistically significant (p=0.11 and p=0.20, respectively). The trials judged at unclear risk of bias (funding) revealed larger effects (p=0.01). Similar to glucosamine, studies with small sample size reported greater effects than large studies (ES difference −0.4, 95% CI −0.7 to −0.1; p=0.01).

### Table 1  Sensitivity analysis of glucosamine for pain at short term

<table>
<thead>
<tr>
<th>Supplement</th>
<th>Variable</th>
<th>Number of studies (patients)</th>
<th>SMD (95% CI)</th>
<th>I², %</th>
<th>Meta-regression</th>
<th>ES difference</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucosamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Randomisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>6 (732)</td>
<td>−0.2 (−0.4 to 0.1)</td>
<td>51</td>
<td></td>
<td>−0.3 (−0.8 to 0.2)</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>4 (314)</td>
<td>−0.6 (−1.3 to 0.1)</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allocation concealment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>3 (525)</td>
<td>0.0 (−0.1 to 0.2)</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>7 (521)</td>
<td>−0.5 (−0.8 to −0.1)</td>
<td>70</td>
<td></td>
<td>−0.4 (−0.9 to 0.0)</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Blinding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>3 (556)</td>
<td>−0.0 (−0.2 to 0.2)</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>6 (437)</td>
<td>−0.4 (−0.9 to −0.0)</td>
<td>76</td>
<td></td>
<td>−0.3 (−0.8 to 0.2)</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>1 (53)</td>
<td>−0.8 (−1.3 to −0.2)</td>
<td>0</td>
<td></td>
<td>−0.7 (−1.6 to 0.1)</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Incomplete outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>6 (839)</td>
<td>−0.1 (−0.3 to 0.1)</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>2 (78)</td>
<td>−1.0 (−1.5 to −0.6)</td>
<td>0</td>
<td></td>
<td>−1.0 (−1.5 to −0.4)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>2 (129)</td>
<td>−0.3 (−0.6 to 0.1)</td>
<td>0</td>
<td></td>
<td>−0.2 (−0.7 to 0.3)</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Selective reporting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>8 (968)</td>
<td>−0.1 (−0.3 to 0.1)</td>
<td>48</td>
<td></td>
<td>−0.9 (−1.5 to −0.4)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>2 (78)</td>
<td>−1.0 (−1.5 to −0.6)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other bias (funding)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>5 (824)</td>
<td>0.0 (−0.1 to 0.2)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>4 (202)</td>
<td>−0.7 (−1.0 to −0.4)</td>
<td>0</td>
<td></td>
<td>−0.7 (−1.0 to −0.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>1 (20)</td>
<td>−1.4 (−2.4 to −0.4)</td>
<td>0</td>
<td></td>
<td>−1.4 (−2.4 to −0.4)</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Study sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>3 (628)</td>
<td>0.1 (−0.1 to 0.2)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>7 (418)</td>
<td>−0.5 (−0.8 to −0.2)</td>
<td>57</td>
<td></td>
<td>−0.5 (−0.9 to −0.2)</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>OA site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee OA</td>
<td>9 (824)</td>
<td>−0.3 (−0.6 to −0.1)</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hip OA</td>
<td>1 (222)</td>
<td>0.1 (−0.2 to 0.3)</td>
<td>0</td>
<td></td>
<td>0.4 (−0.3 to 1.2)</td>
<td>0.28</td>
</tr>
</tbody>
</table>

ES, effect size.; I², measures of heterogeneity; OA, osteoarthritis; SMD, standardised mean difference.

### Table 2  Sensitivity analysis of chondroitin for pain at short term

<table>
<thead>
<tr>
<th>Supplement</th>
<th>Variable</th>
<th>Number of studies (patients)</th>
<th>SMD (95% CI)</th>
<th>I², %</th>
<th>Meta-regression</th>
<th>ES difference</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chondroitin*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Randomisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>5 (1434)</td>
<td>−0.2 (−0.3 to −0.1)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>4 (388)</td>
<td>−0.6 (−0.9 to −0.3)</td>
<td>40</td>
<td></td>
<td>−0.4 (−0.6 to −0.1)</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Allocation concealment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>2 (417)</td>
<td>−0.1 (−0.3 to 0.1)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>7 (1405)</td>
<td>−0.4 (−0.6 to −0.2)</td>
<td>51</td>
<td></td>
<td>−0.3 (−0.6 to 0.1)</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Incomplete outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low risk</td>
<td>7 (1073)</td>
<td>−0.3 (−0.4 to −0.1)</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>2 (749)</td>
<td>−0.6 (−1.2 to −0.1)</td>
<td>83</td>
<td></td>
<td>−0.2 (−0.6 to 0.1)</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Other bias (funding)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>6 (850)</td>
<td>−0.5 (−0.6 to −0.3)</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>3 (972)</td>
<td>−0.2 (−0.3 to −0.0)</td>
<td>0</td>
<td></td>
<td>0.3 (0.1 to 0.6)</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Study sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>5 (1521)</td>
<td>−0.2 (−0.3 to −0.1)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>4 (301)</td>
<td>−0.6 (−0.9 to −0.2)</td>
<td>55</td>
<td></td>
<td>−0.4 (−0.7 to −0.1)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Blinding was all unclear, selective reporting was all low risk, OA site investigated was all knee OA.

ES, effect size.; I², measures of heterogeneity; OA, osteoarthritis; SMD, standardised mean difference.

DISCUSSION

We conducted a comprehensive systematic review on efficacy and safety of various dietary supplements for hand, hip or knee OA including 69 RCTs (11,586 participants) and 20 supplements, the majority of which evaluated glucosamine and chondroitin. Our review demonstrates that the most widely used supplements (eg, glucosamine and chondroitin) are not clinically effective in reducing pain and improving physical function at short term in patients with OA. There were a number of supplements (eg, *Boswellia serrata* extract, pycnogenol, curcumin) with limited studies and participants that were sponsored by pharmaceutical companies that did report moderate to large effects of clinical importance on pain reduction and disability improvement at short term. No data of these supplements on efficacy outcome at medium-term and long-term follow-up were provided by the included trials. No supplements demonstrated clinically important effects for pain at medium term and long term except for green-lipped mussel extract and undenatured type II collagen, which suggested clinically meaningful effects on pain reduction at medium term. Regarding long-term effects on structural improvement among included studies, our review reveals that chondroitin demonstrated small but statistically significant effects when compared with placebo, although with doubtful clinical importance. Overall, our review suggested that with the exception of diacerein, dietary supplements were safe to use.

The strengths of our review include the following: PRISMA recommendations were adhered to, the review was prospectively registered and the GRADE approach was used to appraise the quality of the evidence. We identified a larger number of supplements reported for treating patients with OA than previous reviews, which were limited as narrative reviews and did not perform a meta-analysis. Other reviews only focused on a particular supplement. We included fewer trials of some supplements (eg, chondroitin, glucosamine, diacerein) because we only included RCTs and excluded combination therapy. We included trials of single supplements because they provided the best evidence on the efficacy of specific pharmacological treatments. Different proprietary products of supplements were evaluated in the included studies. We only combined those supplements that had very similar chemical components or different preparations or dosages of same chemicals.

Including more supplements with pooled treatment effects enabled us to provide thorough and valuable information for physicians and/or sports and exercise medicine practitioners to choose suitable supplements within their practice. Further, we pooled analysis for each supplement in categories of different time points as there was a wide range of follow-up durations from 2 weeks to 3 years. To facilitate the interpretation of our results, we have provided clinically interpretable estimates on SMD, and compared our ES with a predetermined smallest worthwhile effect, which reflects the smallest effect of the intervention on outcomes compared with placebo that patients would consider meaningful or important. Moreover, we performed sensitivity analyses for glucosamine and chondroitin to examine the impact of moderator variables that could have influenced our effect estimates, such as risk of bias judgements, sample size and different OA sites by using meta-regression method, and found that small sample size and commercial funding are the main factors that influenced the results, which further contributes to the robustness of our results. However, we are limited by the methodological approach of a standard meta-analysis in which we only compare each of the supplements with placebo; a network meta-analysis might be needed in future to provide a single, coherent ranking of supplements by performing direct and indirect comparisons that allow an individual ranking of each supplement against each other, to determine which is the optimal supplement.

A potential limitation is that we limited the language to English, which could introduce bias. Another limitation is that most of the data pertain to knee OA, which may limit the generalisability to other types of OA. The dosage of supplements was not considered to enable the inclusion of more available trials for each supplement, and we did not consider patient adherence mostly because few studies reported relevant data. The number of studies in each meta-analysis was relatively small for most supplements (only one trial available for eight of them) and long-term data were often unavailable. Nearly 50% of the included trials were at high risk according to the Cochrane risk of bias tool, which highlighted a number of limitations including incomplete outcome data, selective outcome reporting, inadequate description or inappropriate method of randomisation, allocation concealment and blinding. Further, many included studies were sponsored by pharmaceutical companies that often played important roles in some stages of the trial. Importantly, we should be very cautious to explain the GRADE ratings for the treatment effect estimates of single trial (eg, passion fruit peel extract for pain at short term, ES −1.65, GRADE ratings: moderate) because further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

We provided funnel plots for those outcomes with more than 10 studies. The plots clearly show that small study effects/publication bias is likely for the overall analysis, and glucosamine trials in particular, which probably inflates the estimate of efficacy reported. From this point, those supplements with large ES might represent publication bias as well. Furthermore, over two-thirds of the supplements were studied at short term (≤3 months); only one-third of them were studied at medium and long term. We are unsure whether those supplements with large effects at short term would have equivalent efficacy and safety in the long term.

Unanswered questions and future research

Further large and robust RCTs with longer treatment duration are needed to confirm efficacy of supplements with large treatment effects at short term in the presence of limited study quality evidence (passion fruit peel extract, *Curcuma longa* extract, *Boswellia serrata* extract, curcumin, pycnogenol and L-carnitine). It is especially important to further evaluate safety of these supplements in the long run. Further placebo-controlled trials are required to evaluate efficacy of widely used fish oil; we were unable to identify RCTs satisfying the inclusion criteria. Future studies should use appropriate methods to avoid bias in terms of randomisation, allocation concealment, handling dropout rate and avoiding selective outcome reporting. More studies are needed to investigate efficacy and safety of supplements for hand OA. Conflicts of interest had an important influence on the results; future sponsors should avoid playing roles in any stage of any trial.

In summary, supplements are generally safe to use for treating OA. Compared with placebo, the overall analysis including all trials showed that supplements provided moderate and clinical meaningful treatment effects on pain and function in patients with hip, knee and hip OA at short term, although the quality of evidence was very low. The most widely used supplements (eg, glucosamine, chondroitin) do not provide a clinically important effect on knee, hip or hand OA. Some little-known supplements (eg, *Boswellia serrata* extract, pycnogenol, curcumin) appear to have larger effects in the short term, although current data are of low
quality. However, please note that larger effects were seen mostly for supplements that had only one or a few studies and a small number of participants; further research is likely to change the results. Supplements with more numerous triad larger numbers of participants were more likely to have small or non-significant effects. Supplements have no demonstrable clinical effect on pain and function at medium-term and long-term follow-ups.

At this stage, we are unable to recommend the use of diacerein, vitamin D, glucosamine and chondroitin for patients with OA based on the evidence that they do not appear to provide a clinically meaningful benefit. However it is important to note whatever placebo effects may come from use of these supplements. It is worth considering trying *Boswellia serrata* extract, curcumin or pycnogenol with caution over a short period of time (eg, 4–6 weeks) especially for those patients who are very enthusiastic about taking complementary and alternative therapies.

### What is already known?

- Dietary supplements are widely used by patients.
- Clinical guidelines are still controversial.
- Clinical benefits or harms have not been established.

### What are the findings?

- This systematic review and meta-analysis summarises all available randomised placebo-controlled trials on efficacy and safety of dietary supplements to treat hand, hip or knee OA.
- Supplements are found to be generally safe for treating osteoarthritis.
- The most widely used supplements (eg, glucosamine, chondroitin) do not provide a clinically important effect on osteoarthritis.
- Some little-known supplements (eg, *Boswellia serrata* extract, pycnogenol, curcumin) appear to have larger effects, although current data are of low quality.

### Acknowledgements

We would like to thank Jeremy Cullis (Academic Liaison Librarian) for his help with the electronic search. He did not receive additional compensation in association with his support on this article.

### Contributors

XL, GCM, JE, DJH contributed to the design of this study. XL, VR conducted eligible study selection, data extraction and risk of bias assessment. XL, GCM undertook data analysis. XL drafted the manuscript. GCM, JE and DJH edited and reviewed the manuscript.

### Funding

This research was funded by an NHMRC programme grant APP 1091302. XL reports that her time on this review was supported by PuraPharm Postgraduate Scholarship in Integrated Medicine, Sydney Medical School. PuraPharm did not play any role in this review. GCM is supported by an Australian postgraduate award from the Department of Education and Training of Australia. DJH is supported by a NHMRC practitioner fellowship and is a consultant to Flexion, Nestle and Merck Serono.

### Competing interests

None declared.

### Provenance and peer review

Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

### REFERENCES


