Conventional and biologic disease-modifying anti-rheumatic drugs for osteoarthritis: a meta-analysis of randomized controlled trials

Monica S. M. Persson¹,², Aliya Sarmanova¹, Michael Doherty¹,² and Weiya Zhang¹,²

Abstract

Objectives. The role of inflammation in OA is controversial and it is unclear whether suppressing inflammation with conventional or biologic DMARDs is effective. A systematic review and meta-analysis of randomized controlled trials (RCTs) was conducted to compare DMARDs with placebo in participants with symptomatic OA.

Methods. Databases (Medline, Embase, Allied and Complementary Medicine Database, Web of Science and Cochrane Library), conference abstracts and ClinicalTrials.gov were searched to end of November 2017 for placebo-controlled RCTs of DMARDs, including biologics, in symptomatic OA. Pain data at treatment peak time point were extracted and combined using a random-effects meta-analysis. Markers of inflammation and adverse events were extracted and reviewed. Risk of bias assessment was conducted using Cochrane’s tool.

Results. Eleven RCTs (1205 participants) were meta-analysed, including six for conventional DMARDs (757 participants) and five for biologics (448 participants). Overall, DMARDs were statistically superior to placebo [effect size (ES) = 0.18, 95% CI: 0.03, 0.34], although the difference was not clinically significant (0.5 ES threshold). Furthermore, no statistically significant differences were observed in sub-analysis of high-quality trials (ES = 0.11, 95% CI: −0.06, 0.28), biologics (ES = 0.16, 95% CI: −0.02, 0.34) or conventional DMARDs (ES = 0.24, 95% CI: −0.05, 0.54). No difference was found between erosive vs non-erosive hand OA, hand vs knee OA or anti-IL1 vs anti-TNF biologics.

Conclusion. DMARDs did not offer clinically significant pain relief above placebo in OA. This poor efficacy indicates that inflammation may not be a prime driver for OA pain.

Key words: osteoarthritis, DMARDs, biologics, methotrexate, hydroxychloroquine, inflammation, meta-analysis

Introduction

OA is a major cause of pain and disability for which no disease-modifying drug interventions have yet been identified. It is a common complex disorder that involves all joint tissues and affects approximately 1 in 5 women and 1 in 10 men aged >60 years [1]. The pathogenesis of OA has not been fully characterized, and the role of synovial inflammation within this process is intensely debated.

Some perceive OA as the inherent repair process of synovial joints in which modest inflammation is secondary to joint tissue damage caused largely by biomechanical insult [2]. In contrast, others believe synovial inflammation to be a more primary feature and a central driver of OA pain and progression [3]. This belief has encouraged the conduct of randomized controlled trials (RCTs) of conventional and biologic DMARDs in OA. These potent agents are used in inflammatory arthritides, such as RA, where they suppress the aggressive primary inflammation that drives the disease [4].

Conventional and biologic DMARDs are the two major classes of DMARDs used for RA. Conventional DMARDs include drugs such as MTX and HCQ [4], whereas biologic DMARDs are mAbs and soluble receptors that target protein messenger molecules or cells [4].
Whether DMARDs are effective for OA remains controversial. The literature is scattered with hypotheses and hopes for a positive effect in OA. However, this is interspersed with reports of treatment failures. We therefore undertook the present meta-analysis to examine the efficacy of DMARDs, including both conventional and biologic DMARDs, in participants with symptomatic OA.

Methods

Placebo-controlled RCTs comparing a DMARD, including biologics, with placebo in participants with symptomatic OA at any site were included. DMARDs that are recommended or licensed for RA were considered for this review [5, 6]. Supplementary Table S1, available at Rheumatology online, lists all drugs eligible for inclusion in the review. Full text publications and conference abstracts in any language were accepted. No limits were set for publication year.

A systematic literature search was conducted across five databases: Medline, Embase, Allied and Complementary Medicine Database, Web of Science and Cochrane Library. The full search strategies are available in supplementary Table S2, available at Rheumatology online. The search was run from date of database inception to 30 November 2017. Additional trials were searched for using the online clinical trial register ClinicalTrials.gov and EULAR, OARSI and ACR annual meeting abstracts. Following the removal of duplicates, all trials were examined for eligibility. Full texts of eligible abstracts were sought and used. Where full texts were unavailable, conference abstracts were included to minimize publication bias and to ensure the evidence captured was current. The data were extracted independently by two reviewers (M.S.M.P. and A.S.) using a Microsoft Access extraction form created for this review. The following information was extracted: publication details, including author, journal, year and publication type (full text or abstract); trial details, including trial funder, study design, blinding and duration; participant details and demographics, including number of participants, age, gender distributions and BMI; the joint affected; method of diagnosis (e.g. clinical, radiographic); OA subset details; and intervention/control details, including the drug, formulation, dose and frequency.

Pain data at treatment peak time point were extracted. Treatment peak time point (i.e. time point where treatment group had the greatest improvement) was chosen under the assumption that if a difference between treatment and placebo was not observed at this time point, one could confidently assume no efficacy at other time points. Where multiple tools for assessing pain were presented, the outcome was chosen using the hierarchy defined by Jüni et al. [7]. Where possible, changes from baseline pain scores were extracted/calculated; if not, end point scores were used to calculate the between-group mean differences. If pain outcomes were dichotomized, the percentage of participants with improvement in pain (as defined in the publication) was used. Data were preferentially extracted from intention-to-treat analyses. Data were not extracted from graphs, and missing data were not sought from investigators. Where a trial examined multiple dosages of the intervention, these were combined into one group for analyses.

Other outcomes extracted were inflammation (local or systemic) and incidence of adverse events (AEs).

Risk of bias

Two reviewers independently assessed the risk of bias of included trials using Cochrane’s Risk of Bias Tool [8]. The distribution of scores for each domain was presented.

Statistical methods

Hedges’ effect size (ES) and corresponding 95% CI were calculated. If dichotomized, the odds ratio for improvement was extracted/calculated and converted to an ES [8, 9]. The study estimates were combined using a random-effects meta-analysis weighted using inverse-variance methods. Heterogeneity was quantified using $I^2$ and the $P$-value was calculated using the $Q$ test [10, 11]. Publication bias was presented using a funnel plot, and the asymmetry of this plot was examined by Egger’s test [12]. As recommended in the National Institute for Health and Care Excellence guidelines for OA, a minimal clinically important difference threshold of ES = 0.5 for the lower level of the 95% CI was used to determine clinical significance [13].

All DMARDs were pooled before being examined by type (conventional vs biologic). Further subgroup analysis was conducted to examine biologics by the mechanism of action (TNF-inhibitor vs IL1-inhibitor). Additional subgroup analyses were conducted by joint affected, OA subtype (erosive vs non-erosive hand OA) and publication type. A sensitivity analysis was only conducted for high-quality trials using adequate allocation concealment as an indicator of quality [14].

Analyses were done with Stata (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX, USA: StataCorp LP). The trial was registered with PROSPERO (PROSPERO 2017 CRD42017067427).

Results

Description of trials

Thirteen studies were identified comparing a DMARD with placebo in participants with OA (Fig. 1 and Table 1). Outcome data were not available for extraction from two abstracts [15, 16]. The meta-analysis is based on the remaining 11 RCTs (7 full texts, 4 abstracts), including 6 for conventional [25–27] and 5 for biologic [18–20, 24, 23] DMARDs. A variety of DMARDs were examined, including HCQ (5 trials) [17, 21, 22, 26, 27], MTX (1 trial) [25], anakinra (ANK; 1 trial) [19], adalimumab (ADA, 3 trials) [18, 20, 24] and etanercept (ETN; 1 trial) [23]. Ten trials were parallel design trials, while one trial [18] was a cross-over design trial and combined both treatment periods. Median trial duration was 24 weeks (range 12–52 weeks).

A total of 1205 participants with clinically and radiographically confirmed OA were included in the meta-analysis. Five trials examined participants with knee OA, of which one [17] was restricted to participants with clinical
signs indicative of synovitis. In contrast, the RCT by Chevalier et al. [19] excluded participants with effusions and inflammatory flares. The remaining six trials studied participants with hand OA. Three of these trials [18, 23, 24] were restricted to participants with erosive hand OA. In addition, a further hand OA trial was limited to participants with OA that was refractory to analgesics [20].

In the six studies that presented the gender distributions of participants, approximately three-quarters of the participants were women (78.2%). The mean age of participants in the eight trials where age data were available ranged from 47.9 years [21] to 62.6 years [19].

**Bias within and across trials**

Potential risk of bias was demonstrated in the analysed trials (Fig. 2). The primary sources of bias were selective outcome reporting and incomplete outcome data, where over half the studies were associated with high risks of bias. Online trial registrations were found for six studies, and a published protocol was found for one study. These were compared with the outcomes available in the included publication, and the publications most commonly did not include all pre-specified outcomes, did not fully present the outcomes or conducted analyses that had not been pre-specified. Furthermore, many trials excluded participants from analysis [17, 18, 22, 23], used inappropriate methods for imputation of data (e.g. last observation carried forward) [19, 24], or did not provide sufficient detail regarding the amount of missing data or how it was handled [20, 21, 25, 26, 27]. These domains cause concerns about the risk of bias of the included trials.

All trials were reported as double-blinded; however, the adequacy of the blinding methods could not be determined in 5 of the 11 trials, due to insufficient information in the publications. The remaining RCTs described appropriate procedures for blinding of participants, physicians and outcome assessors.

Nearly half (45%) of the studies reported adequate methods of allocation concealment, most commonly through central allocation external to the study investigators. One trial [25] allocated treatment by the order of enrolment to the trial, and this was deemed to have a high risk of bias. There was no evidence of funnel plot asymmetry (P = 0.121) (supplementary Fig. S1, available at Rheumatology online).
### Table 1 Description of trials comparing DMARD with placebo in OA

<table>
<thead>
<tr>
<th>References</th>
<th>Funding source</th>
<th>Design</th>
<th>Drug</th>
<th>Route</th>
<th>Dose</th>
<th>Study duration (week)</th>
<th>Duration analysed (week)</th>
<th>No. randomized (n on DMARD)</th>
<th>No. of female (%)</th>
<th>Mean age (s.d.)</th>
<th>OA Joint detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abou-Raya et al. [16]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>MTX</td>
<td>Oral</td>
<td>25 mg/week + usual treatment</td>
<td>16</td>
<td>Not analysed</td>
<td>88 (44)</td>
<td>NP</td>
<td>NP</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Abou-Raya et al. [17]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>HCQ</td>
<td>Oral</td>
<td>400 mg/day</td>
<td>36</td>
<td>16</td>
<td>166 (83)</td>
<td>NP</td>
<td>NP</td>
<td>Knee OA + clinical signs of synovitis</td>
</tr>
<tr>
<td>Aitken et al. [18]</td>
<td>Unclear</td>
<td>Cross-over</td>
<td>ADA</td>
<td>SC</td>
<td>40 mg/every other week</td>
<td>12</td>
<td>12</td>
<td>43 (43)</td>
<td>NP</td>
<td>61 (8.4)</td>
<td>Hand OA + synovitis (erosive hand OA)</td>
</tr>
<tr>
<td>Chevalier et al. [19]</td>
<td>Commercial</td>
<td>Parallel</td>
<td>ANK</td>
<td>IA</td>
<td>50 mg or 150 mg (single injection)</td>
<td>12</td>
<td>4</td>
<td>170 (101)</td>
<td>107 (63)</td>
<td>62.6 (9.7)</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Chevalier et al. [20]</td>
<td>Commercial</td>
<td>Parallel</td>
<td>ADA</td>
<td>SC</td>
<td>40 mg (two injections separated by 2-week interval)</td>
<td>26</td>
<td>6</td>
<td>85 (42)</td>
<td>71 (84)</td>
<td>62.5 (6.9)</td>
<td>Hand OA refractory to analgesics</td>
</tr>
<tr>
<td>Jokar et al. [21]</td>
<td>Public</td>
<td>Parallel</td>
<td>HCQ</td>
<td>Oral</td>
<td>400 mg/day</td>
<td>24</td>
<td>21</td>
<td>51 (25)</td>
<td>43 (98)</td>
<td>47.9 (9.8)</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Kingsbury et al. [22]</td>
<td>Public</td>
<td>Parallel</td>
<td>HCQ</td>
<td>Oral</td>
<td>Dose dependant on gender, height and weight</td>
<td>52</td>
<td>52</td>
<td>248 (114)</td>
<td>NP</td>
<td>NP</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Kloppenburg et al. [23]</td>
<td>Commercial</td>
<td>Parallel</td>
<td>ETA</td>
<td>SC</td>
<td>50 mg/week (first 24 weeks), thereafter 25 mg/week</td>
<td>52</td>
<td>52</td>
<td>90 (45)</td>
<td>NP</td>
<td>60</td>
<td>Hand OA + synovitis (erosive hand OA)</td>
</tr>
<tr>
<td>Lindsley et al. [15]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>INF</td>
<td>IA</td>
<td>100 mg (single injection)</td>
<td>8</td>
<td>Not analysed</td>
<td>16 (8)</td>
<td>NP</td>
<td>NP</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Verbruggen et al. [24]</td>
<td>Commercial</td>
<td>Parallel</td>
<td>ADA</td>
<td>SC</td>
<td>40 mg/every other week</td>
<td>52</td>
<td>52</td>
<td>60 (30)</td>
<td>NP</td>
<td>61.3 (6.5)</td>
<td>Hand OA + synovitis (erosive hand OA)</td>
</tr>
<tr>
<td>Holanda et al. [25]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>MTX</td>
<td>Oral</td>
<td>7.5 mg/week</td>
<td>17</td>
<td>17</td>
<td>58 (29)</td>
<td>48 (83)</td>
<td>61.1 (8.8)</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Bonfante et al. [26]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>HCQ</td>
<td>Oral</td>
<td>400 mg/day</td>
<td>16</td>
<td>16</td>
<td>32 (16)</td>
<td>26 (81)</td>
<td>60.7 (9.6)</td>
<td>Knee OA</td>
</tr>
<tr>
<td>Lee et al. [27]</td>
<td>Unclear</td>
<td>Parallel</td>
<td>HCQ</td>
<td>Oral</td>
<td>400 mg/day</td>
<td>24</td>
<td>24</td>
<td>196 (98)</td>
<td>168 (86)</td>
<td>58 (7.6)</td>
<td>Hand OA</td>
</tr>
</tbody>
</table>

ADA: adalimumab; ANK: anakinra; ETA: etanercept; IA: intraarticular; INF: infliximab; NP: not presented.
Efficacy of DMARDs in OA

Pain

Overall, the pooled result of outcomes for conventional and biologic DMARDs was statistically superior to that for outcomes for placebo with respect to pain relief (Fig. 3 and Table 2). The ES was 0.18 (95% CI: 0.03, 0.34). The estimate was associated with a moderate level of inconsistency ($I^2 = 41.7\%$) [8]; however, the test for heterogeneity was not statistically significant ($P = 0.071$). A sensitivity analysis including only high-quality trials, using adequate allocation concealment as a quality indicator, showed no difference between DMARDs and placebo (ES = 0.11, 95% CI: −0.06, 0.28) (Table 2). No heterogeneity was evident in the estimate. Furthermore, separate examination of conventional and biologic DMARDs found that neither conventional (ES = 0.24, 95% CI: −0.05, 0.54) nor biologic (ES = 0.16, 95% CI: −0.02, 0.34) DMARDs were superior to placebo for pain relief. Large and significant heterogeneity was observed across the conventional DMARD trials. However, the results for biologic DMARDs were homogeneous, and this homogeneity was retained when examined by mechanism of action. Neither IL1-inhibitors (ES = 0.14, 95% CI: −0.16, 0.45) nor TNF-inhibitors (ES = 0.17, 95% CI: −0.05, 0.54) were effective.

Further subgroup analyses found that the lack of efficacy of treatment compared with placebo did not vary by the joint site (hand vs knee OA), inflammatory phenotype (erosive vs non-erosive hand OA) or publication type (conference abstract or full text).

Inflammation and AEs

Only six studies assessed inflammation. One found a statistically significant improvement (US synovitis) [16], one found a clinically relevant improvement (clinical and US) [17], and four found no difference between treatment and placebo (MRI, serum and tissue CRP, effusion) [15, 18, 20, 24]. However, no quantitative data were available for further analysis. AE incidence was poorly reported across trials and was not aggregated.

Discussion

This first meta-analysis of conventional and biologic DMARDs in OA, including 11 RCTs with over 1200 participants, did not demonstrate statistically significant pain relief from either conventional or biologic DMARDs compared with placebo. Biologics, in particular, showed homogeneity across studies. Conventional and biologic DMARDs, when combined, were statistically superior to placebo (ES = 0.18, 95% CI: 0.03, 0.34), but this was far below the minimal clinically important difference threshold (ES = 0.5) used in the UK [13]. Moreover, examination of only high-quality trials with adequate allocation concealment did not produce statistically significant results (ES = 0.11, 95% CI: −0.06, 0.28), and this estimate was homogeneous. Any benefit observed is likely overestimated by the risk of bias associated with selective outcome reporting [8]. Furthermore, the analysis is based on peak time point for the intervention, so even at their most effective time-point these treatments do little better than placebo.

Although the use of biologic DMARDs in OA has been heralded as a promising way forward based on in vitro, animal model and uncontrolled pilot studies, narrative reviews have indicated that they do not appear to have any clear benefit over placebo in RCTs [28–30]. This is confirmed by the present meta-analysis, which is able to provide quantitative evidence for a lack of statistical superiority of biological DMARDs over placebo. In contrast, the perceived efficacy of conventional DMARDs in the literature is mixed. For example, the ACR specifically advises against the use of MTX and does not offer any recommendations on HCQ due to limited evidence [31], whereas others favour a positive effect for MTX and HCQ.
FIG. 3 Forest plot of DMARDs vs placebo in OA

Hedges’ effect size (ES) and 95% CI presented for pain at peak time point for intervention arm. ADA: adalimumab; ANK: anakinra; ETN: etanercept.

TABLE 2 Overall, sensitivity and subgroup meta-analysis results for biologic and conventional DMARDs in OA

<table>
<thead>
<tr>
<th>Analysis</th>
<th>No. of trials</th>
<th>No. analysed</th>
<th>ES (95% CI)</th>
<th>$I^2$ (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>11</td>
<td>1205</td>
<td>0.18 (0.03 to 0.34)</td>
<td>41.7% (0.071)</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>5</td>
<td>765</td>
<td>0.11 (--0.06, 0.28)</td>
<td>0.0% (0.818)</td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMARD type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologic DMARD</td>
<td>5</td>
<td>448</td>
<td>0.16 (--0.02, 0.34)</td>
<td>0.0% (0.975)</td>
</tr>
<tr>
<td>IL1-inhibitors</td>
<td>1</td>
<td>170</td>
<td>0.14 (--0.16, 0.45)</td>
<td></td>
</tr>
<tr>
<td>TNF-inhibitors</td>
<td>4</td>
<td>278</td>
<td>0.17 (--0.05, 0.39)</td>
<td>0.0% (0.926)</td>
</tr>
<tr>
<td>Conventional DMARD</td>
<td>6</td>
<td>757</td>
<td>0.24 (--0.05, 0.54)</td>
<td>70.0% (0.005)</td>
</tr>
<tr>
<td>Joint</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knee</td>
<td>5</td>
<td>477</td>
<td>0.34 (--0.05,0.73)</td>
<td>71.7% (0.007)</td>
</tr>
<tr>
<td>Hand</td>
<td>6</td>
<td>728</td>
<td>0.09 (--0.05, 0.24)</td>
<td>0.0% (0.948)</td>
</tr>
<tr>
<td>OA type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erosive hand OA</td>
<td>3</td>
<td>193</td>
<td>0.19 (--0.06, 0.45)</td>
<td>0.0% (0.846)</td>
</tr>
<tr>
<td>Non-erosive hand OA</td>
<td>3</td>
<td>535</td>
<td>0.05 (--0.12, 0.22)</td>
<td>0.0% (0.979)</td>
</tr>
<tr>
<td>Publication type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conference abstract</td>
<td>4</td>
<td>547</td>
<td>0.11 (--0.05, 0.27)</td>
<td>0.0% (0.925)</td>
</tr>
<tr>
<td>Full text</td>
<td>7</td>
<td>658</td>
<td>0.27 (--0.01, 0.55)</td>
<td>62.7% (0.013)</td>
</tr>
</tbody>
</table>

Presented as Hedges’ effect size (ES) and associated 95% CI. $I^2$ (the variation in ES attributable to heterogeneity) and associated $P$-values are presented.
DMARDS were unlikely to be effective. It also allowed more trials to be analysed. Displaying only a small effect at their peak, conventional and biologic DMARDs are unlikely to have considerable, or in fact any, effect at other time points.

Conclusion

Neither conventional nor biological DMARDs relieve pain due to OA. There is no difference between anti-IL1 and anti-TNF biologics, and no difference between treating erosive vs non-erosive hand OA or hand vs knee OA. Although combining all DMARDs provided statistically significant pain relief, this was not clinically significant, nor was it supported by sensitivity analyses of high-quality trials. The results suggest that inflammation may not be a principal risk factor for OA pain.

Funding: This work was supported by Arthritis Research UK [grant number 20777].

Disclosure statement: M.D. reports a grant from AstraZeneca funding a non-drug PI-led study in Nottingham (Sons of Gout Study) and honoraria for Advisory Boards on osteoarthritis and gout for AstraZeneca, Grunenthal, Mallinckrodt and Roche, outside the submitted work. W.Z. reports honoraria for Grunenthal and speaker fees for Bioiberica and Hisun, outside the submitted work. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at Rheumatology online.

References


