Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic disease modifying antirheumatic drugs for rheumatoid arthritis: abridged Cochrane systematic review and network meta-analysis

Glen S Hazlewood,[1, 2, 3, 4] Cheryl Barnabe,[1, 2, 4] George Tomlinson,[5] Deborah Marshall,[2, 4] Dan Devoe,[4] Claire Bombardier[5, 6, 7]

For numbered affiliations see end of article.

Correspondence to: G S Hazlewood, 3330 Hospital Dr NW, Calgary, AB, Canada T2N1N1 glenhazlewood@gmail.com

ABSTRACT

OBJECTIVE
To compare methotrexate based disease modifying antirheumatic drug (DMARD) treatments for rheumatoid arthritis in patients naive to or with an inadequate response to methotrexate.

DESIGN
Systematic review and Bayesian random effects network meta-analysis of trials assessing methotrexate used alone or in combination with other conventional synthetic DMARDs, biologic drugs, or tofacitinib in adult patients with rheumatoid arthritis.

DATA SOURCES
Trials were identified from Medline, Embase, and Central databases from inception to 19 January 2016; abstracts from two major rheumatology meetings from 2009 to 2015; two trial registers; and hand searches of Cochrane reviews.

STUDY SELECTION CRITERIA
Randomized or quasi-randomized trials that compared methotrexate with any other DMARD or combination of DMARDs and contributed to the network of evidence between the treatments of interest.

MAIN OUTCOMES
American College of Rheumatology (ACR) 50 response (major clinical improvement), radiographic progression, and withdrawals due to adverse events. A comparison between two treatments was considered statistically significant if its credible interval excluded the null effect, indicating >97.5% probability that one treatment was superior.

RESULTS
158 trials were included, with between 10 and 53 trials available for each outcome. In methotrexate naive patients, several treatments were statistically superior to oral methotrexate for ACR50 response: sulfasalazine and hydroxychloroquine ("triple therapy"), several biologics (abatacept, adalimumab, etanercept, infliximab, rituximab, tocilizumab), and tofacitinib. The estimated probability of ACR50 response was similar between these treatments (range 56-67%), compared with 41% with methotrexate. Methotrexate combined with adalimumab, etanercept, certolizumab, or infliximab was statistically superior to oral methotrexate for inhibiting radiographic progression, but the estimated mean change over one year with all treatments was less than the minimal clinically important difference of 5 units on the Sharp-van der Heijde scale. Triple therapy had statistically fewer withdrawals due to adverse events than methotrexate plus infliximab. After an inadequate response to methotrexate, several treatments were statistically superior to oral methotrexate for ACR50 response: triple therapy, methotrexate plus hydroxychloroquine, methotrexate plus leflunomide, methotrexate plus intramuscular gold, methotrexate plus most biologics, and methotrexate plus tofacitinib. The probability of response was 61% with triple therapy and ranged widely (27-70%) with other treatments. No treatment was statistically superior to oral methotrexate for inhibiting radiographic progression. Methotrexate plus abatacept had a statistically lower rate of withdrawals due to adverse events than several treatments.

CONCLUSIONS
Triple therapy (methotrexate plus sulfasalazine plus hydroxychloroquine) and most regimens combining biologic DMARDs with methotrexate were effective in controlling disease activity, and all were generally well tolerated in both methotrexate naive and methotrexate exposed patients.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Meta-analyses have shown that most biologic disease modifying antirheumatic drugs (DMARDs) combined with methotrexate are superior to methotrexate alone for controlling disease activity

However, the benefits of combining methotrexate with conventional synthetic DMARDs are uncertain

A Cochrane network meta-analysis of biologic treatments for rheumatoid arthritis showed some differences but did not compare biologic therapy against combination therapy with conventional synthetic DMARDs

Understanding the comparative benefits and harms of these treatments is essential, given that biologic therapy costs 10-20 times more than most conventional synthetic DMARDs

WHAT THIS STUDY ADDS

This Cochrane network meta-analysis included 158 trials with more than 37 000 patients

"Triple therapy" (methotrexate plus sulfasalazine plus hydroxychloroquine) was superior to methotrexate alone and not statistically different from methotrexate plus any biologic DMARD or tofacitinib for controlling disease activity, either as initial therapy or after an inadequate response to methotrexate

Given the low cost of triple therapy relative to biologic DMARDs and tofacitinib, these findings support a therapeutic trial of triple therapy as initial treatment or after an inadequate response to methotrexate

Introduction
Methotrexate based treatments form the core of rheumatoid arthritis treatment. Methotrexate is

recommended as the first disease modifying antirheumatic drug (DMARD) for most patients with rheumatoid arthritis, and methotrexate co-prescription is recommended when using biologic DMARDs or the recently approved tofacitinib.[12] Combining methotrexate with other conventional synthetic DMARDs, however, is more controversial. A trial of conventional synthetic DMARD combination therapy before biologic therapy is not recommended by either major rheumatology guideline, although each provides the option.[12] Understanding the comparative benefits and harms of these treatments is essential to inform decision making, as treatment with biologic DMARDs or tofacitinib costs 10-20 times more than treatment with methotrexate and most conventional synthetic DMARDs.

Network (mixed treatment) meta-analyses are a natural avenue of comparative effectiveness research, as they combine all direct and indirect evidence to estimate treatment effects between all treatments of interest.[3] If treatments A and B are in the same study, direct evidence links A and B. If they are compared in separate studies with a common comparator C, then the A-C and B-C studies allow an indirect comparison of A and B. Longer chains of indirect comparisons (A-B, B-C, C-D) are also possible. Considering indirect evidence is critical if a treatment decision must be made and the treatments have not been directly compared in a head-to-head trial. Indirect evidence is also important to consider when head-to-head trials are available, as it adds to the entire body of evidence and may help to refine the precision in estimation of the treatment effect.[3]

A previous Cochrane network meta-analysis examined the relative effects of different biologic therapies through indirect comparisons and found some differences between agents.[4] Our review expands on this, by including combination therapy with methotrexate plus conventional synthetic DMARDs. A previous traditional (non-network) Cochrane meta-analysis did not find an additional overall benefit with combination therapy over methotrexate alone.[5] By including indirect evidence, we expand the evidence base available to draw from. For example, three recently published trials compared combination therapy with methotrexate plus sulfasalazine plus hydroxychloroquine versus methotrexate plus anti-tumor necrosis factor (TNF) therapy.[6-8] The inclusion of these trials in a network meta-analysis adds indirect evidence on the relative effects of methotrexate plus sulfasalazine plus hydroxychloroquine compared with all other treatments in the network.

Methods

This is an abridged version of a Cochrane systematic review. The protocol and Cochrane review (once published) can be accessed from the Cochrane Library.[9]

Inclusion criteria

Study type

We included randomized trials or quasi-randomized trials (where allocation to treatment groups was not truly random—for example, alternate patients) of at least 12 weeks' duration that contained any intervention of interest (defined in detail below) and could be linked within the network through a shared comparator. For example, if we identified a trial comparing methotrexate with hydroxychloroquine, the trial was included if another trial was available that compared hydroxychloroquine with any other treatment of interest. We divided trials into groups based on patients' previous exposure to methotrexate: either methotrexate naive or inadequate response to methotrexate. We excluded trials that required all patients to have failed to respond to anti-TNF therapy.

Participants

We included studies of adults (age >18 years) with rheumatoid arthritis, according to 1958, 1987, or 2010 classification criteria.[10-12]

Interventions

The interventions of interest were oral methotrexate; parenteral (intramuscular or subcutaneous) methotrexate; methotrexate in combination with conventional synthetic DMARD therapy including antimalarials (hydroxychloroquine/chloroquine), sulfasalazine, leflunomide, ciclosporin, intramuscular gold, and azathioprine; methotrexate in combination with biologic DMARD therapy (adalimumab, certolizumab pegol, etanercept, golimumab, infliximab, abatacept, rituximab, tocilizumab); and methotrexate in combination with tofacitinib. We applied no dose restriction to conventional synthetic DMARDs, given the variability of dosing in clinical practice. Biologic DMARDs and tofacitinib were limited to currently recommended doses or dose equivalents.

Outcomes

The major outcomes of the review were American College of Rheumatology (ACR) 50 response, a composite measure of improvement in disease activity (dichotomous outcome);[13] radiographic progression, measured by Larsen, Sharp, or modified Larsen/Sharp scores (continuous outcome);[14] and withdrawals due to adverse events, including death (dichotomous outcome). Multiple secondary outcomes were evaluated and are reported in the full Cochrane review.[9]

Search methods

We did an electronic database search in Medline, Embase, and Cochrane Central from inception to 19 January 2016. The search strategies contained subject headings and keywords for "rheumatoid arthritis", "methotrexate", and "randomized controlled trial". The search strategy is available in the full Cochrane review.[9] We also searched the trial registries ClinicalTrials.gov (http://clinicaltrials.gov/) and the International Clinical Trials Registry Platform (http://apps.who.int/trialsearch/) by using the search term "rheumatoid arthritis AND methotrexate". Finally, we hand searched for abstracts from American College of Rheumatology and European League Against Rheumatism conferences from 2009 to 2015 and reviewed existing Cochrane reviews. All languages were included.

Study selection

Two review authors (GH, ChB) independently screened articles for inclusion by title or abstract and full text if necessary. Disagreements were resolved by consensus or discussion with a third review author (ClB).

Data extraction and handling of missing data

Three review authors working in pairs (GH, ChB, DD) abstracted relevant data from included studies into an Excel spreadsheet. Characteristics of trials and baseline characteristics of patients were extracted by one author (GH) and confirmed by a second (ChB or DD); outcome data were extracted independently, with disagreements resolved through discussion. For all trials, we also sought data from clinical trial registries (clinicaltrials.gov, www.who.int/ictrp/en/), the US Food and Drug Administration, the European Medicines Agency, and drug manufacturers' websites. For continuous measures, if standard deviations were not available, we calculated them from available variance data (for example, standard errors) if possible. For toxicity outcomes, if the drug exposure was not available, we calculated it; we assumed withdrawals to occur at a constant rate, unless specific information was available to permit a more accurate calculation. If data were presented only in graphical format, two independent reviewers (GH, ChB) extracted numbers digitally and averaged or corrected them if an obvious discrepancy was apparent.

Time point of outcome assessment

For parallel group trials, we used end of trial data for all outcomes. For crossover trials or trials comparing different DMARD sequences, we used data at the time of the initial switch in treatment for efficacy outcomes (ACR50 response and radiographic progression). For withdrawals due to adverse events in crossover and strategy trials, we abstracted data from the longest follow-up available, provided the adverse event was assigned to the treatment received at the time of the adverse event (on-treatment data). In certain trials, a switch in treatment was allowed or required for patients who failed to achieve a certain response at a given time (usually 12-16 weeks). Patients who receive early escape therapy are typically treated as treatment failures at all future time points. For early escape trials, we used end of trial (carried forward) data for ACR50 and radiographic progression. For withdrawals due to adverse events, we used on-treatment data from the longest follow-up available, similar to the approach for crossover trials.

Risk of bias of studies

The methodological quality of included trials was independently assessed using the Cochrane Collaboration's tool for assessing risk of bias by two review authors or by one author (GH) if his assessment agreed with a published Cochrane review. Studies were graded as having a "low risk," "high risk," or "unclear risk" of bias across the seven specified domains.[15] The risk of bias was assessed separately for the three major outcomes for domains in which the risk of bias could differ. The domains "blinding of participants" and "blinding of outcome assessment" were assessed separately for radiographic progression; "incomplete outcome data" was assessed separately for each of the three major outcomes. For each of the three major outcomes, we also judged an overall risk of bias.

Data analysis

For the primary analyses, we excluded trials with a high risk of bias for that outcome. We evaluated treatment effects by using odds ratios for ACR50 response and a standardized mean difference for radiographic progression. We estimated standardized mean differences by dividing the modeled change in each arm by the within trial pooled standard deviation of the change value. We summarized withdrawals due to adverse events as rate ratios to allow for differences in exposure between arms in early escape and crossover trials.

We fitted random effects bayesian network meta-analyses for each outcome measure.[16-18] The outcome measures were standardized mean differences for the continuous outcome radiographic progression, logarithms of odds ratios for the dichotomous outcome ACR50 response, and logarithms of rate ratios for withdrawals due to adverse events. The model code was based on published work and is presented in annotated form in appendix A.[16-18] In this model, each trial of a pair of treatments is assumed to estimate a particular treatment effect, which varies around a mean effect with a shared between study variance (in fixed effects models in our sensitivity analyses, the between study variance around the mean effect was set to zero). This mean treatment effect is then broken down into "basic parameters" that are unique for each treatment. The basic parameter was set to "0" (no effect) for oral methotrexate, so that the basic parameter for the other treatments provided the treatment effect relative to oral methotrexate. From these basic parameters, we determined the treatment effect between every pair of treatments. This model assumes consistency between the indirect and direct evidence, such that both contribute to estimation of the same treatment effect.

We used uninformative prior probability distributions for all parameters. We used 10 000 burn-in iterations followed by 10 000 monitoring iterations. We assessed convergence by running three chains, inspecting the sampling history plots, and calculating Gelman-Rubin-Brooks statistics.[19] We assessed model fit by using residual deviance and the deviance information criterion. We used R statistical software version 3.1.2 (www.r-project.org) for all data analyses, with rjags package version 3-14 running Just Another Gibbs Sampler (JAGS) version 3.4.0.[20]

Presentation of results

We report the posterior median and 95% credible interval for all pairwise treatment effects. To facilitate interpretation of the results, we considered an effect to be "significant" if its 95% credible interval excluded the null effect, which equates to a 97.5% probability that one of the treatments is superior to the other. We recognized, however, that this cut-off was arbitrary and

therefore also calculated the probability that each treatment was superior to each other and the rank ordering of treatments.

We converted the average treatment effect for each outcome into an absolute response by using an assumed (baseline) value for oral methotrexate. For all analyses, the assumed baseline value was the median from a bayesian random effects model of the oral methotrexate arms. For ACR50 response, we used all trials in the analysis to estimate the assumed probability of response. For radiographic progression, we calculated the assumed mean over one year on the Sharp-van der Heijde scale from the trials that reported this outcome for oral methotrexate. We then calculated the absolute effect for each treatment by using this assumed value and the mean differences for each treatment relative to oral methotrexate on the Sharp-van der Heijde scale, which we calculated by multiplying the standardized mean differences by the pooled within arm standard deviation for studies that used the Sharp-van der Heijde scale. For withdrawals due to adverse events, we estimated the assumed rate at one year from the available trials and converted it to an absolute probability by using the rate ratio, assuming that the time to withdrawal over one year was exponentially distributed for each person.

Quality of evidence (GRADE)

We used GRADE (Grading of Recommendations Assessment, Development and Evaluation) guidance to assess the quality of evidence from a network meta-analysis, with the quality of evidence graded from high (best) to very low (worst).[21] This approach considers the quality of both the direct and indirect evidence, as well as the consistency (coherence) of the indirect and direct evidence and likelihood of "intransitivity," which exists if heterogeneity is present in the trials that form the different comparisons within the indirect evidence. We did "node splitting" to separate the indirect evidence from the direct evidence to inform these evaluations.[22] We calculated a statistical measure of inconsistency by comparing the treatment effects from the indirect and direct evidence where both were available and calculating the probability that they were the same; lower values indicate a higher likelihood of inconsistency.[22]

Meta-regression and sensitivity analyses

We did meta-regression for several pre-specified trial level characteristics. Selected post-hoc sensitivity analyses were added, including fixed effect models for the major outcome "radiographic progression," as few trials were available to estimate a random effect. Full details of these analyses are provided in the full review.

We also did sensitivity analyses around several modeling assumptions. In the protocol, we had planned to use odds ratios to pool withdrawals due to adverse events. However, we changed the analyses to rate ratios given the differences in exposure between arms in early escape and crossover trials. As a sensitivity analysis, we compared the rate ratios with odds ratios, in which we used the total exposure (in patient months) in each arm

as the denominator, instead of the number of patients. The model estimates the effect on the monthly odds of an outcome, assuming independence between months, and should approximate the rate ratio from a Poisson model.

The choice of prior distribution for the between study variance may affect the estimated treatment effects, although this effect has been found to be small in analyses of 10 or more studies.[23] For the primary analysis, we followed published guidance and chose a prior that was vague but realistic.[23] We then did sensitivity analyses using an additional uninformative prior and potentially informative priors of Turner et al (odds ratio for ACR50 response and rate ratio for withdrawals due to adverse events)[24] and Rhodes et al (radiographic progression).[25]

Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in the dissemination of results.

Results

Search results and description of included studies

From 9817 unique records, we identified 158 trials including more than 37 000 patients meeting our inclusion criteria (fig 1). Methotrexate dosing varied across studies and was variably reported (table 1; full list of studies in appendix B). Most trials enrolled patients with high disease activity (median baseline swollen joint count 15.1), with a similar distribution across drug classes (table 1).

In the methotrexate naive network (fig 2, top), most comparisons of methotrexate plus biologic DMARDs were against methotrexate, with no head-to-head comparisons between different biologic DMARDs. Trials evaluating conventional synthetic DMARD therapy were generally smaller but more interconnected than trials of biologic DMARDs. In the methotrexate inadequate response network (fig 2, bottom), connections between methotrexate plus biologic DMARDs and methotrexate were large in size (more patients), whereas connections between conventional synthetic DMARDs were few and small in size. Four head-to-head trials of biologic therapy formed links between several biologic therapies,[26-29] and all four trials that compared methotrexate plus biologic therapy against methotrexate plus conventional synthetic DMARDs were included in this network.[6-8 30] The network diagrams show all trials; the actual numbers of trials for each outcome varied and are reported below.

Methodological quality of included studies

The risk of bias of the trials varied considerably across each domain (fig 3). The overall risk of bias was high in 30% of trials for ACR50 response, in 21% for radiographic progression, and in 17% for withdrawals due to adverse events. These trials were excluded from the primary analysis.
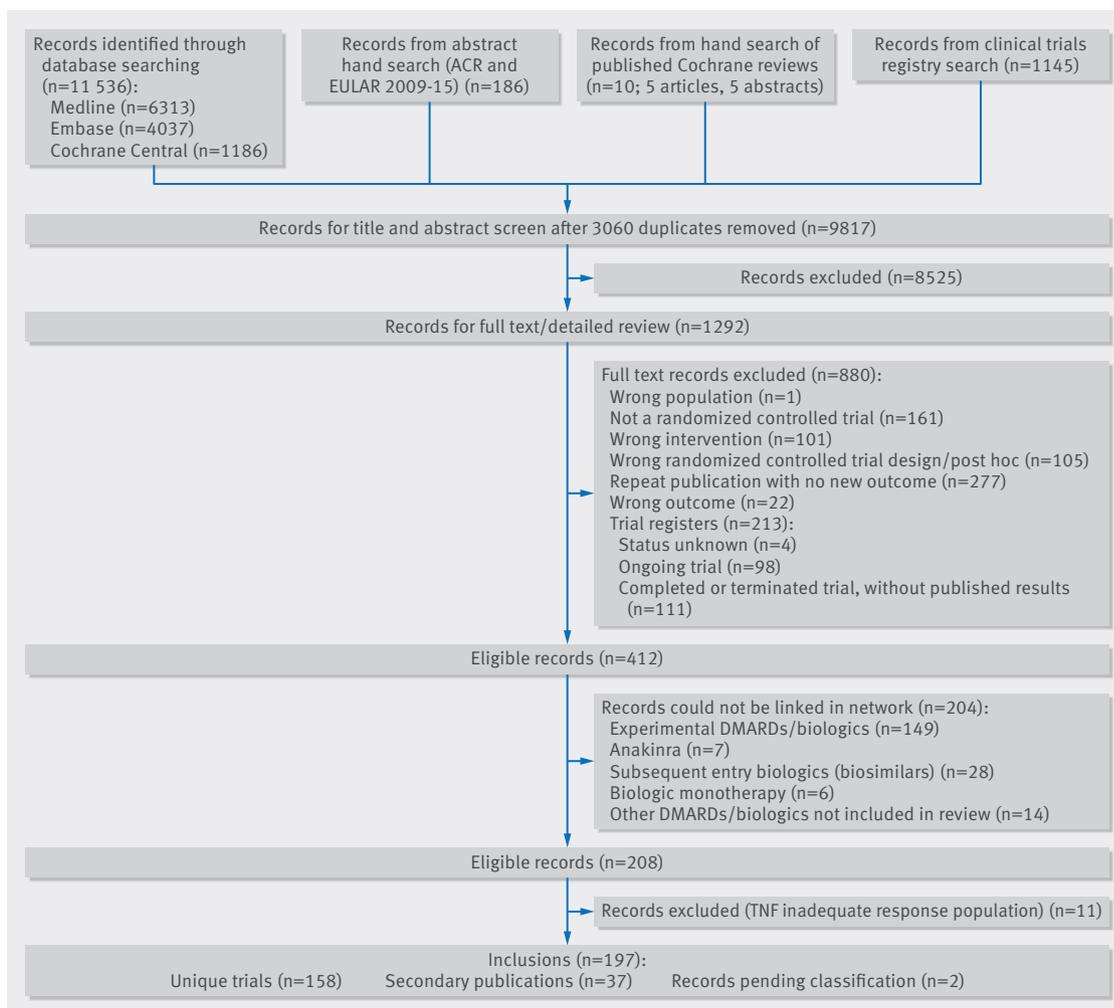
Fig 1 | Flow chart of search

Effects of interventions: major outcomes

Tables 2 and 3 summarize the main results; the pairwise results for each comparison and ranking of treatments are in appendix C. Appendix D shows details of the appraisal of the quality of evidence using the GRADE approach.

Methotrexate naive patients

ACR50—Twenty nine trials with 10 697 patients were included in this analysis. The combination of methotrexate plus several biologic DMARDs (intravenous abatacept, adalimumab, etanercept, infliximab, rituximab, tocilizumab 8 mg/kg) and methotrexate plus tofacitinib were statistically superior to oral methotrexate, with the 95% credible interval excluding the null effect (table 2). A high probability existed that methotrexate plus subcutaneous abatacept and methotrexate plus tocilizumab 4 mg/kg were also superior to oral methotrexate (97% for both) (table 2). "Triple therapy" (methotrexate plus sulfasalazine plus hydroxychloroquine) was the only conventional synthetic DMARD combination that had a statistically significant higher odds of ACR50 response than oral methotrexate. This comparison was based on indirect evidence and was

judged to be moderate quality. The magnitude of the estimated probability of ACR50 response was similar between triple therapy (61.2%, 95% credible interval 44.2 to 76.5) and the other DMARDs that had a statistically significant benefit relative to oral methotrexate (point estimate range 56-67%). In comparison, the estimated probability of ACR50 response with oral methotrexate was 40.5%. In pairwise comparisons, we found no statistically significant difference between triple therapy and methotrexate plus any biologic DMARD, although we could not rule out an important difference, as the credible intervals were wide for some comparisons (appendix C).

Radiographic progression—Eighteen trials with 7594 patients were included in this analysis. The combinations of methotrexate plus several biologic DMARDs (adalimumab, certolizumab, etanercept, infliximab) were associated with a statistically significant reduction in radiographic progression relative to oral methotrexate (table 2). The probability of superiority to oral methotrexate was more than 95% with methotrexate plus rituximab and tocilizumab 8 mg/kg, with the 95% credible interval narrowly including the null effect. The sizes of the effects for all interventions relative to oral

Table 1 | Summary of trial characteristics

Drug	No of studies	No of patients	Year published, median (range)	Early escape design, % of studies (No of patients)	Trial duration*, median (range) weeks	MTX dose >15 mg/wk†, % of studies (No of patients)	Disease duration, median (range) years	Swollen joint count, median (range)
MTX + bDMARDs/tofacitinib:								
MTX + etanercept	10	2448	2007 (1999-2014)	0	38 (12-52)	50 (n=1833)	2 (0.5-13)	13.9 (11-24)
MTX + infliximab	13	2806	2006 (2000-14)	8 (n=264)	26 (13-54)	46 (n=1990)	7.6 (0.4-10.5)	15 (5-21.5)
MTX +adalimumab	16	4465	2010 (2003-15)	38 (n=1936)	24 (12-104)	50 (n=1809)	2.5 (0.1-11.7)	16.3 (8.7-22.5)
MTX + rituximab	4	1262	2008 (2004-12)	25 (n=342)	24 (24-104)	50 (n=683)	8.6 (0.9-11.5)	20.9 (20.2-21.6)
MTX + abatacept	10	3612	2012 (2005-15)	0	25 (17-52)	60 (n=3014)	6.4 (0.5-9.3)	17.1 (10-22.4)
MTX + tocilizumab	10	4859	2012 (2006-16)	50 (n=3671)	24 (16-52)	60 (n=2729)	6.6 (0.4-9.2)	13.7 (6.4-20.1)
MTX + certolizumab	7	2680	2012 (2008-15)	71 (n=1561)	24 (24-52)	29 (n=1119)	6 (0.3-9.6)	21 (16.4-22.5)
MTX + golimumab	6	1640	2012 (2008-13)	83 (n=1570)	24 (16-52)	50 (n=1132)	6.9 (3.2-8.7)	13.5 (11.6-15.4)
MTX + tofacitinib	4	749	2012 (2011-15)	50 (n=621)	24 (12-52)	50 (n=621)	8.7 (0.7-9.1)	14.7 (14.1-14.9)
Subtotal	80	24521	2011 (1999-2016)	31 (n=9965)	24 (12-104)	50 (n=14930)	6.3 (0.1-13)	16.4 (5-24)
Comparative effectiveness trials:								
Head-to-head bDMARDs/tofacitinib	4	1658	2010 (2006-14)	25 (n=501)	27 (26-104)	50 (n=1077)	7.8 (1.8-11.3)	16.6 (15.9-20.6)
MTX + bDMARDs v MTX + csDMARDs	4	1382	2012 (2012-13)	0	63 (16-104)	25 (n=353)	0.5 (0.3-5.2)	12 (11.2-12.8)
Subtotal	8	3040	2012 (2006-14)	12 (n=501)	27 (16-104)	38 (n=1430)	5.2 (0.3-11.3)	15.9 (11.2-20.6)
MTX + csDMARDs:								
MTX + azathioprine	1	209	1992	0	24	0	8.6	17.3
MTX + hydroxychloroquine/ chloroquine	7	452	2005 (1993-2012)	0	26 (12-52)	0	4 (0.3-7.7)	10.7 (8.2-27.3)
MTX + sulfasalazine	6	639	2000 (1994-2007)	0	52 (24-76)	0	0.4 (0.2-5)	16.7 (9.8-22.6)
MTX + ciclosporin	9	1100	2003 (1995-2008)	11 (n=120)	48 (16-104)	22 (n=64)	1.1 (0.2-10.3)	13.6 (11-19)
MTX + sulfasalazine + hydroxychloroquine	4	503	2005 (1996-2013)	0	91 (13-104)	0	4.4 (0.5-8.6)	17.1 (9.5-29.8)
MTX + leflunomide	3	921	2006 (2002-15)	0	24 (16-36)	67 (n=455)	6.2 (0.7-11.6)	14.3 (10.7-18)
MTX + IM gold	1	65	2005	0	48	100 (n=65)	3.2	11
Subtotal	31	3889	2003 (1992-2015)	3 (n=120)	48 (12-104)	16 (n=584)	1.1 (0.2-11.6)	13.6 (8.2-29.8)
MTX v csDMARD monotherapy:								
Placebo	5	324	1985 (1985-93)	20 (n=52)	14 (12-18)	0	9.5 (4.8-14)	27.5 (24-30.9)
Azathioprine	5	257	1991 (1987-94)	0	24 (24-52)	0	12 (8.7-13.9)	14.6 (9.5-21.9)
IM gold	5	489	1991 (1988-2001)	0	48 (26-52)	20 (n=99)	4 (1.2-6)	14 (13.9-15.2)
Sulfasalazine	2	211	1998 (1995-2002)	0	24	0	4 (1.4-6.6)	9.3
Ciclosporin	2	203	1999 (1998-2000)	0	65 (26-104)	50 (n=100)	3.8 (2.2-5.5)	13.1 (12.2-14)
Leflunomide	16	3258	2002 (1999-2014)	12 (n=567)	24 (12-52)	25 (n=927)	3.9 (0.5-6.8)	11.8 (8.2-16.5)
Hydroxychloroquine	2	409	2006 (2000-12)	0	24	0	1.5 (1-2.1)	NA
SC v oral MTX	2	467	2009 (2008-10)	50 (n=383)	24	100 (n=467)	0.2	15
Subtotal	39	5618	2000 (1985-2014)	10 (n=1002)	24 (12-104)	21 (n=1593)	4.5 (0.2-14)	14 (8.2-30.9)
Total	158	37068	2008 (1985-2016)	20 (n=11588)	24 (12-104)	35 (n=18537)	4.8 (0.1-14)	15.1 (5-30.9)

Trials are grouped by comparator and sorted by year of first trial within each class, for illustrative purposes. Patients' characteristics, including number of patients, include only arms considered in review.
DMARD=disease modifying antirheumatic drug; bDMARD; biologic DMARD; csDMARD=conventional synthetic DMARD; tsDMARD=targeted synthetic DMARD; IM=intramuscular; MTX=methotrexate; SC=subcutaneous; SJC=swollen joint count.
*Trial duration for efficacy outcomes; some studies had longer follow-up for safety outcomes.
†Studies for which dose of methotrexate could be confirmed as ≥15 mg/week; in some studies, methotrexate was dosed across a range of values that included 15 mg/week, but the actual dose was not provided.

methotrexate were small. The expected radiographic progression was 2.34 points over one year with oral methotrexate (the reference treatment) and lower for all other treatments, which is below the minimal clinically important difference of 5 units on the Sharp-van der Heijde scale.[31] We found no statistically significant differences between treatments in pairwise comparisons (appendix C).

In post-hoc sensitivity analyses using fixed effects models, the point estimates were nearly identical to the random effects model but the credible intervals were not as wide, resulting in several biologic DMARDs (plus methotrexate) for which the treatment effects reached statistical significance relative to oral methotrexate (appendix C). Methotrexate plus sulfasalazine plus hydroxychloroquine

was not statistically superior to oral methotrexate in either the random effects or fixed effect models.

Withdrawals due to adverse events—Thirty seven trials with a total follow-up of 10 528 patient years were included in this analysis. The combination of methotrexate plus azathioprine had a statistically significant increase in withdrawals due to adverse events compared with oral methotrexate and several other treatments (table 2 and appendix C). A high probability also existed that intramuscular/subcutaneous methotrexate plus ciclosporin, methotrexate plus infliximab, or methotrexate plus tocilizumab 8 mg/kg had a higher rate of withdrawals due to adverse events than oral methotrexate (97%, 97%, and 95%), with the results narrowly failing to reach statistical significance.
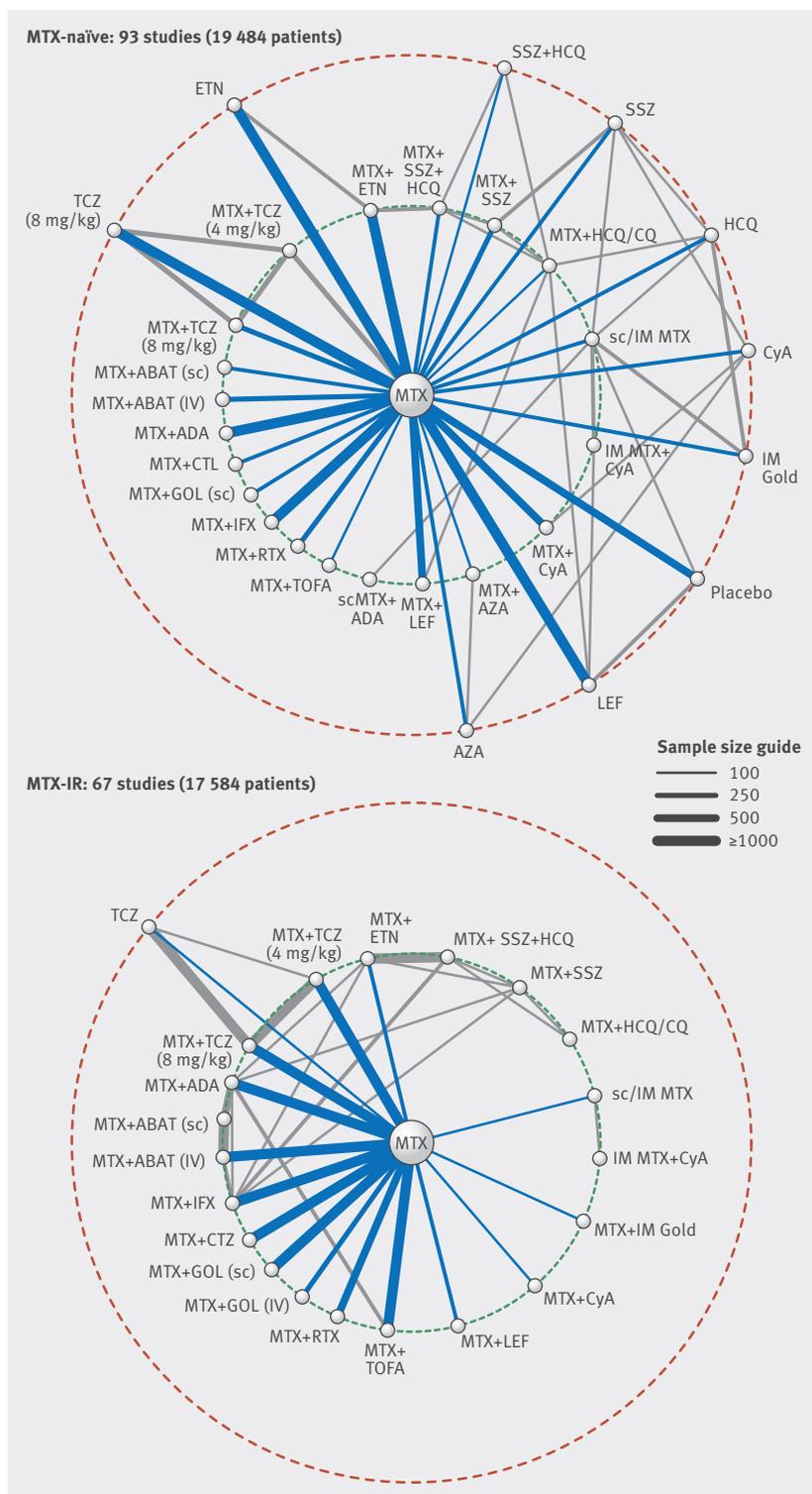
Fig 2 | Networks of included studies for methotrexate naive (top) and methotrexate inadequate response populations (bottom). Each line represents direct comparison between two treatments from one or more trials. Biologic/targeted synthetic DMARDs are shown on left of each network and conventional synthetic DMARDs on right. Treatments on innermost circle (green hashed line) are treatments of interest, whereas treatments on outermost circle (red hashed line) are other treatments that form links between treatments of interest. Comparisons against methotrexate are shown in blue. Two trials were included in both analyses. ABAT=abatacept; ADA=adalimumab; AZA=azathioprine; CQ=chloroquine; ETN=etanercept; CTL-certolizumab; CyA=ciclosporin; GOL=golimumab; HCQ=hydroxychloroquine; IFX=infliximab; IM=intramuscular; IR=inadequate response; IV=intravenous; LEF=leflunomide; MTX=methotrexate; RTX=rituximab; sc=subcutaneous; SSZ=sulfasalazine; TOFA=tofacitinib; TCZ=tocilizumab

We found no statistically significant differences in pairwise comparisons between different biologic DMARDs (plus methotrexate). Methotrexate plus sulfasalazine plus hydroxychloroquine was associated with a statistically lower rate of withdrawals due to adverse events than methotrexate plus infliximab (rate ratio 0.26, 95% credible interval 0.06 to 0.91).

Methotrexate inadequate response patients
ACR50—Forty five trials with 12 549 patients were included in this analysis. Several treatments were superior to oral methotrexate for ACR50 response (table 3). The results reached statistical significance for the combination of methotrexate and several conventional synthetic DMARDs (sulfasalazine plus hydroxychloroquine, hydroxychloroquine, leflunomide, or intramuscular gold), methotrexate plus all biologic DMARDs with available evidence, and methotrexate plus tofacitinib. The estimated probability of an ACR50 response with triple therapy was 60.5% (39.4% to 81.8%) and varied widely for other treatments (point estimate range 27-70%). We found no evidence for certolizumab, as the available trials were judged to be at high risk of bias. In general, the credible intervals in the pairwise comparisons between different treatments combinations were wide, although some estimates reached statistical significance (appendix C): methotrexate plus etanercept was superior to the combination of methotrexate plus most biologic DMARDs, and methotrexate plus sulfasalazine plus hydroxychloroquine was superior to methotrexate plus the biologic DMARDs intravenous abatacept, infliximab, and tocilizumab 4 mg/kg.

The quality of the evidence for triple therapy (methotrexate plus sulfasalazine plus hydroxychloroquine) versus methotrexate was judged to be moderate, as some minor inconsistencies existed in the findings of the two trials that compared triple therapy with methotrexate plus etanercept,[6 8] and because the study design of one of the trials was judged to indirectly address the comparison of interest.[8] This trial randomized patients at baseline to a step-up to triple therapy versus a step-up to methotrexate plus etanercept only if an inadequate response to methotrexate was found after six months.[8]

Radiographic progression—Ten trials with 3238 patients were included in this analysis. We found no statistically significant differences between any treatment and oral methotrexate, although the probability of superiority ranged from 76% (methotrexate plus subcutaneous golimumab) to 94% (methotrexate plus infliximab) (table 3). Similar to the analysis in methotrexate naive patients, the credible intervals were more precise in the post-hoc fixed effect model, resulting in several treatments that reached statistical significance relative to oral methotrexate (methotrexate plus abatacept (intravenous and subcutaneous), adalimumab, etanercept, intravenous golimumab, and infliximab) (appendix C). A 96% probability existed that methotrexate plus sulfasalazine plus hydroxychloroquine was superior to oral methotrexate in the fixed effects model (standardized mean difference −0.40, 95% credible interval −0.84 to 0.04).
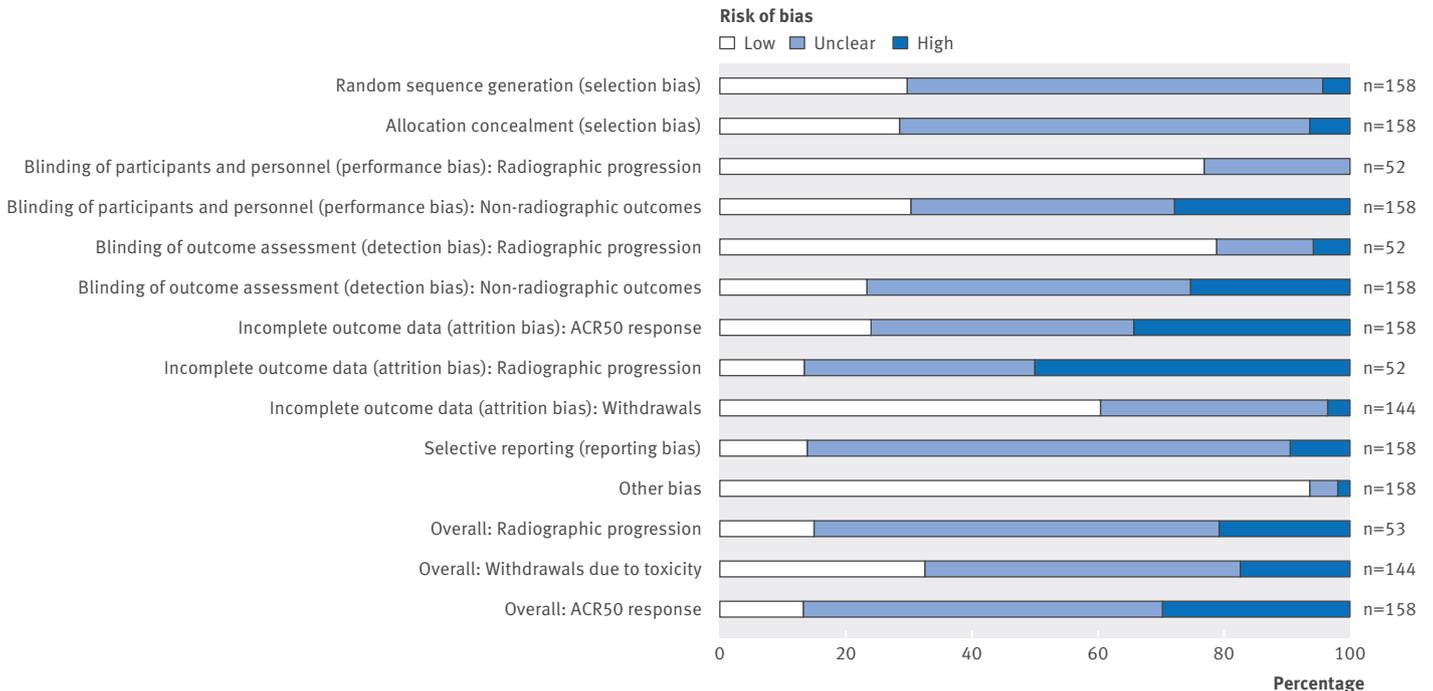
Fig 3 | Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies. ACR50=American College of Rheumatology 50 response

Withdrawals due to adverse events—Fifty three trials with a total follow-up of 9950 patient years were included in this analysis. Methotrexate plus ciclosporin and methotrexate plus tocilizumab 8 mg/kg were the only treatments with statistically significant higher rates of withdrawals due to adverse events relative to oral methotrexate (table 3). In pairwise comparisons, methotrexate plus subcutaneous abatacept and methotrexate plus intravenous abatacept were associated with a statistically significant lower rate of withdrawals due to adverse events than several treatments, including methotrexate plus biologic DMARDs and methotrexate plus sulfasalazine plus hydroxychloroquine (appendix C).

GRADE quality assessment

Appendix D shows the overall quality of evidence for each outcome. The assessment of consistency in the evidence was limited, as few comparisons had both indirect and direct evidence available. The only comparison that was downgraded for inconsistency was the comparison of tocilizumab 8 mg/kg versus methotrexate in the analysis of ACR50 response in patients with inadequate response to methotrexate. The odds ratio for the direct evidence was 1.68 (0.62 to 4.56), compared with 4.15 (1.72 to 9.63), with a P value of 0.17. We did not detect evidence of publication bias, although this was difficult to assess, as we had too few trials for any direct comparison to permit formal tests of funnel plot asymmetry.

Minor outcomes

Results for the minor outcomes are reported in detail in the full Cochrane review. Compared with oral methotrexate, some treatments had a statistically significant

higher rate of certain adverse events: methotrexate plus sulfasalazine plus hydroxychloroquine and methotrexate plus sulfasalazine had a higher rate of total gastrointestinal events (excluding oral and liver toxicity) in the analysis of methotrexate naive patients (rate ratio 2.10 (95% credible interval 1.19 to 3.96) and 1.90 (1.18 to 2.99)), methotrexate plus leflunomide had a higher rate of alanine aminotransferase elevations in the analysis of patients with inadequate response to methotrexate (4.75, 1.16 to 20.70), and methotrexate plus tocilizumab 8 mg/kg had a higher rate of leukopenia in this analysis (16.25, 1.48 to 206).

Additional analyses

Meta-regression

In meta-regression analyses, we found a significant association between the odds ratio for ACR50 response and certain study level covariates (appendix C), but the adjusted treatment effects were largely unchanged (appendix C). When all studies (both methotrexate naive and methotrexate inadequate response patients) were included in the same network meta-analysis and the network assignment was specified with a meta-regression covariate, the odds ratios of methotrexate inadequate response trials were 2.05 (1.70 to 2.48) times higher.

Sensitivity analyses for ACR50 response

When we excluded studies with partial methotrexate exposure from the methotrexate naive analysis, the comparison of triple therapy against oral methotrexate for ACR50 response was no longer statistically significant (appendix C). The point estimate, however, was slightly higher than in the main analysis (favoring triple

Table 2 | Summary of findings: methotrexate naive patients

Intervention	Absolute risk (95% CrI) No of events/1000 patients at 1 year	Average treatment effect relative to oral MTX (95% CrI) Odds ratio	Probability treatment superior to oral MTX %	No of trials providing direct evidence	Quality of evidence
ACR50 (29 studies; 10697 patients)					
MTX	405	Reference	–	–	–
MTX + abatacept (IV)	555 (407 to 699)	1.84 (1.01 to 3.42)	98	1	High
MTX + abatacept (SC)	574 (390 to 730)	1.98 (0.94 to 3.97)	97	1	High
MTX + adalimumab	588 (508 to 661)	2.10 (1.52 to 2.87)	>99	4	High
IM/SC MTX + adalimumab	601 (353 to 805)	2.22 (0.80 to 6.06)	94	0	Moderate (imprecision)
MTX + certolizumab	504 (361 to 646)	1.49 (0.83 to 2.68)	93	1	Moderate (study limitations)
MTX + etanercept	671 (578 to 757)	3.00 (2.02 to 4.59)	>99	2	High
MTX + golimumab (SC)	476 (315 to 638)	1.33 (0.68 to 2.59)	83	1	Moderate (study limitations)
MTX + infliximab	580 (470 to 719)	2.03 (1.30 to 3.77)	>99	3	High
MTX + rituximab	622 (469 to 750)	2.42 (1.30 to 4.42)	99	1	High
MTX + tocilizumab (4 mg/kg)	529 (392 to 665)	1.66 (0.95 to 2.92)	97	1	Moderate (study limitations)
MTX + tocilizumab (8 mg/kg)	565 (426 to 696)	1.91 (1.09 to 3.36)	98	2	High
MTX + tofacitinib	674 (416 to 864)	3.04 (1.05 to 9.37)	98	1	Moderate (imprecision)
MTX + ciclosporin	539 (370 to 695)	1.72 (0.86 to 3.36)	94	1	Low (indirectness, imprecision, study limitations)
IM/SC MTX + ciclosporin	516 (234 to 803)	1.57 (0.44 to 6.01)	75	0	Low (imprecision, study limitations)
MTX + hydroxychloroquine/chloroquine	346 (136 to 663)	0.78 (0.23 to 2.90)	35	0	Moderate (imprecision)
MTX + sulfasalazine	427 (219 to 654)	1.10 (0.41 to 2.78)	57	1	Low (indirectness, imprecision, study limitations)
MTX + sulfasalazine + hydroxychloroquine	612 (442 to 765)	2.32 (1.17 to 4.79)	99	0	Moderate (indirectness)
IM/SC MTX	434 (288 to 595)	1.13 (0.59 to 2.16)	65	1	Moderate (study limitations)
Radiographic progression (18 studies; 7594 patients)	Mean change on Sharp-VdH scale over 1 year (points)	Standardized mean difference	%		
MTX	2.34	Reference	–	–	–
MTX + abatacept (IV)	1.11 (−1.29 to 3.47)	−0.20 (−0.60 to 0.19)	88	1	Moderate (imprecision)
MTX + adalimumab	0.09 (−1.52 to 1.88)	−0.37 (−0.64 to −0.08)	99	2	High
MTX + certolizumab	−0.01 (−1.74 to 1.74)	−0.39 (−0.68 to −0.10)	99	2	Moderate (study limitations)
MTX + etanercept	0.12 (−1.19 to 1.67)	−0.37 (−0.59 to −0.11)	99	3	High
MTX + golimumab (SC)	1.57 (−0.87 to 4.08)	−0.13 (−0.53 to 0.29)	76	1	Low (study limitations, imprecision)
MTX + infliximab	−0.26 (−2.59 to 2.10)	−0.43 (−0.82 to −0.04)	98	1	High
MTX + rituximab	0.03 (−2.40 to 2.42)	−0.38 (−0.79 to 0.01)	97	1	Moderate (imprecision)
MTX + tocilizumab (4 mg/kg)	0.84 (−1.64 to 3.30)	−0.25 (−0.66 to 0.16)	91	1	Moderate (study limitations)
MTX + tocilizumab (8 mg/kg)	0.14 (−2.28 to 2.54)	−0.37 (−0.77 to 0.03)	97	1	Moderate (study limitations)
MTX + tofacitinib	1.09 (−2.78 to 5.17)	−0.21 (−0.85 to 0.47)	73	1	Moderate (imprecision)
MTX + ciclosporin	1.07 (−0.68 to 2.94)	−0.21 (−0.50 to 0.10)	92	2	Low (study limitations, imprecision)
MTX + sulfasalazine + hydroxychloroquine	2.14 (−2.18 to 6.69)	−0.03 (−0.75 to 0.72)	54	0	Moderate (imprecision)

(Continued)

Table 2 | Summary of findings: methotrexate naive patients

Intervention	Absolute risk (95% CrI) No of events/1000 patients in 1 year	Average treatment effect relative to oral MTX (95% CrI) Rate ratio	Probability treatment superior to oral MTX %	No of trials providing direct evidence	Quality of evidence
Withdrawals due to adverse events (37 studies; 10 528 patient years)					
MTX	76	Reference	–	–	–
MTX + abatacept (IV)	52 (15 to 163)	0.70 (0.21 to 2.35)	74	1	High
MTX + abatacept (SC)	71 (15 to 310)	0.97 (0.20 to 4.89)	52	1	Moderate (imprecision)
MTX + adalimumab	88 (46 to 153)	1.21 (0.63 to 2.18)	24	4	High
IM/SC MTX + adalimumab	60 (5.1 to 458)	0.81 (0.07 to 8.06)	58	0	Moderate (imprecision)
MTX + etanercept	59 (33 to 117)	0.80 (0.45 to 1.64)	77	3	High
MTX + golimumab (SC)	164 (49 to 520)	2.36 (0.67 to 9.67)	8	1	Low (study limitations, imprecision)
MTX + infliximab	175 (69 to 448)	2.53 (0.94 to 7.81)	3	4	Moderate (imprecision)
MTX + rituximab	61 (17 to 204)	0.83 (0.22 to 3.01)	62	1	High
MTX + tocilizumab (4 mg/kg)	96 (35 to 249)	1.33 (0.46 to 3.77)	25	1	Low (study limitations, imprecision)
MTX + tocilizumab (8 mg/kg)	158 (61 to 384)	2.26 (0.82 to 6.38)	5	1	Low (study limitations, imprecision)
MTX + tofacitinib	66 (13 to 293)	0.90 (0.17 to 4.56)	55	1	Moderate (imprecision)
MTX + azathioprine	356 (113 to 842)	5.79 (1.58 to 24.31)	1	3	Moderate (indirectness)
MTX + ciclosporin	77 (28 to 166)	1.06 (0.37 to 2.38)	44	2	Moderate (indirectness)
IM/SC MTX + ciclosporin	491 (71 to 999)	8.89 (0.98 to 139.30)	3	0	Very low (extreme imprecision, indirectness)
MTX + hydroxychloroquine/chloroquine	98 (30 to 392)	1.35 (0.40 to 5.26)	30	1	Low (imprecision)
MTX + sulfasalazine	95 (49 to 190)	1.31 (0.67 to 2.78)	21	4	Moderate (indirectness)
MTX + sulfasalazine + hydroxychloroquine	49 (21 to 109)	0.67 (0.28 to 1.51)	84	2	Moderate (imprecision)
IM/SC MTX	131 (42 to 399)	1.85 (0.56 to 6.69)	14	1	Moderate (imprecision)

CrI=credible interval; IM=intramuscular; IV=intravenous; MTX=methotrexate; SC=subcutaneous; VdH=van der Heijde.

therapy) and higher than the comparison of any other treatment versus methotrexate. When we included trials at high risk of bias in the methotrexate inadequate response ACR50 response analysis, methotrexate plus certolizumab and subcutaneous/intramuscular methotrexate had a statistically significant higher odds of ACR50 response compared with oral methotrexate (appendix C). Little change in the point estimates for ACR50 response occurred at different time points of assessment, although the credible intervals were wider for several comparisons (appendix C). This supports the meta-regression results in which no association was found between trial duration and ACR50 response.

Modeling assumptions

The rate ratios for withdrawals due to adverse events were similar to the odds ratios for both the methotrexate naive and methotrexate inadequate response analyses (appendix E, tables E1 and E2). For results with wide credible intervals, the point estimates for the rate ratios and odds ratios showed a larger difference, as expected, but this would not change the interpretation of the results. The choice of prior distribution around the between study variance had little effect on the posterior distribution of the between study variance (appendix E, table E3). The results of ACR50 response were very similar for the alternative uninformative prior and slightly more precise for the informative prior, but this would also not affect the interpretation of our results (appendix E, tables E4 and E5).

Discussion

Our systematic review and network meta-analysis compared methotrexate and all currently used combinations of DMARDs with methotrexate. The main new finding from our review was that methotrexate plus sulfasalazine plus hydroxychloroquine ("triple therapy") was superior to oral methotrexate for ACR50 response, in both methotrexate naive and methotrexate inadequate response populations. We found a statistically significant benefit for other conventional synthetic DMARD combinations compared with oral methotrexate, but only after an inadequate response to methotrexate, and the magnitude of effect or quality of evidence was graded lower than for triple therapy. Most biologic DMARDs, in combination with methotrexate, had a statistically significant benefit compared with oral methotrexate for ACR50 response in both methotrexate naive and methotrexate inadequate response populations. Importantly, the magnitude of effect for ACR50 response was higher for triple therapy than for most biologic DMARDs, and we did not find any statistical benefit for methotrexate plus biologic therapy compared with triple therapy. This has important policy implications given the difference in cost between these treatments. Methotrexate combined with adalimumab, certolizumab, etanercept, or infliximab had a statistically significant benefit for inhibiting joint damage compared with oral methotrexate, but the effect was small and observed only in methotrexate naive patients. Most treatments were well tolerated, although some

Table 3 | Summary of findings: methotrexate inadequate response patients

Intervention	Absolute risk (95% CrI)	Average treatment effect relative to oral MTX (95% CrI)	Probability treatment superior to oral MTX	No of trials providing direct evidence	Quality of evidence
ACR50 (45 studies; 12 549 patients)	No of events/1000 patients at 1 year	Odds ratio	%		
MTX	127	Reference	–	–	–
MTX + abatacept (IV)	357 (290 to 437)	3.81 (2.80 to 5.33)	>99	5	High
MTX + abatacept (SC)	377 (284 to 488)	4.16 (2.72 to 6.53)	>99	0	High
MTX + adalimumab	389 (330 to 462)	4.37 (3.38 to 5.89)	>99	10	High
MTX + etanercept	642 (456 to 818)	12.31 (5.76 to 30.78)	>99	3	Moderate (study limitations)
MTX + golimumab (SC)	395 (273 to 539)	4.49 (2.57 to 8.01)	>99	3	Moderate (study limitations, indirectness)
MTX + golimumab (IV)	343 (207 to 514)	3.58 (1.79 to 7.25)	>99	1	Moderate (study limitations)
MTX + infliximab	335 (264 to 422)	3.46 (2.46 to 5.00)	>99	6	High
MTX + rituximab	343 (241 to 477)	3.59 (2.18 to 6.27)	>99	3	High
MTX + tocilizumab (4 mg/kg)	273 (171 to 399)	2.57 (1.42 to 4.56)	>99	2	High
MTX + tocilizumab (8 mg/kg)	377 (264 to 499)	4.16 (2.46 to 6.85)	>99	3	High
MTX + tofacitinib	441 (325 to 568)	5.42 (3.31 to 9.01)	>99	3	High
MTX + hydroxychloroquine/chloroquine	566 (241 to 871)	8.94 (2.18 to 46.14)	>99	0	Low (high imprecision)
MTX + IM gold	704 (228 to 988)	16.34 (2.03 to 553.42)	>99	1	Very low (extreme imprecision)
MTX + leflunomide	453 (245 to 703)	5.69 (2.23 to 16.27)	>99	1	Moderate (imprecision)
MTX + sulfasalazine	267 (67 to 667)	2.50 (0.49 to 13.76)	87	0	Low (high imprecision)
MTX + sulfasalazine + hydroxychloroquine	605 (394 to 818)	10.51 (4.46 to 30.81)	>99	0	Moderate (study limitations)
Radiographic progression (10 studies; 3238 patients)	Mean change on Sharp-VdH scale over 1 year (points)	Standardized mean difference	%		
MTX	3.35	Reference	–	–	–
MTX + abatacept (IV)	1.45 (−5.85 to 8.80)	−0.30 (−1.44 to 0.85)	84	1	Moderate (imprecision)
MTX + abatacept (SC)	0.26 (−9.65 to 11.10)	−0.48 (−2.03 to 1.21)	86	0	Moderate (imprecision)
MTX + adalimumab	0.51 (−6.42 to 7.96)	−0.44 (−1.53 to 0.72)	90	1	Moderate (study limitations, imprecision)
MTX + etanercept	−0.49 (−12.09 to 11.06)	−0.60 (−2.41 to 1.21)	87	0	Moderate (imprecision)
MTX + golimumab (SC)	2.44 (−2.77 to 7.66)	−0.14 (−0.96 to 0.67)	76	2	Low (study limitations, inconsistency, indirectness, imprecision)
MTX + golimumab (IV)	0.52 (−6.56 to 7.98)	−0.44 (−1.55 to 0.73)	89	1	Low (study limitations, imprecision)
MTX + infliximab	−1.08 (−8.34 to 6.35)	−0.69 (−1.83 to 0.47)	94	1	Low (study limitations, imprecision)
MTX + sulfasalazine + hydroxychloroquine	0.70 (−9.58 to 11.05)	−0.41 (−2.02 to 1.20)	82	0	Low (indirectness, imprecision)
Withdrawals due to adverse events (53 studies; 9950 patient years)	No of events/1000 patients in 1 year	Rate ratio	%		
MTX	73	Reference	–	–	–
MTX + abatacept (IV)	54 (31 to 90)	0.76 (0.44 to 1.30)	85	6	High
MTX + abatacept (SC)	39 (21 to 72)	0.55 (0.28 to 1.03)	97	0	Moderate (indirectness)
MTX + adalimumab	100 (67 to 155)	1.44 (0.95 to 2.30)	4	8	High
MTX + certolizumab	99 (56 to 196)	1.42 (0.79 to 2.99)	13	4	Low (study limitations, indirectness)
MTX + etanercept	89 (40 to 195)	1.28 (0.56 to 2.92)	29	3	Moderate (study limitations)
MTX + golimumab (SC)	72 (28 to 184)	1.02 (0.39 to 2.78)	48	3	Low (study limitations, indirectness)
MTX + golimumab (IV)	92 (26 to 370)	1.32 (0.36 to 6.31)	34	1	Low (study limitations, imprecision)
MTX + infliximab	112 (70 to 179)	1.62 (0.99 to 2.70)	3	6	High
MTX + rituximab	141 (53 to 376)	2.07 (0.74 to 6.45)	8	3	Moderate (imprecision)
MTX + tocilizumab (4 mg/kg)	112 (67 to 191)	1.63 (0.95 to 2.90)	4	3	High
MTX + tocilizumab (8 mg/kg)	118 (74 to 188)	1.71 (1.01 to 2.84)	2	5	High
MTX + tofacitinib	87 (52 to 152)	1.24 (0.74 to 2.26)	21	4	High
MTX + ciclosporin	212 (84 to 503)	3.27 (1.20 to 9.57)	1	2	Low (indirectness, imprecision)
MTX + IM gold	260 (35 to 999)	4.12 (0.49 to 102.75)	10	1	Very low (extreme imprecision)
MTX + leflunomide	127 (53 to 290)	1.86 (0.74 to 4.68)	8	1	Moderate (imprecision)
MTX + sulfasalazine + hydroxychloroquine	125 (62 to 249)	1.82 (0.87 to 3.92)	5	0	Moderate (imprecision)

CrI=credible interval; IM=intramuscular; IV=intravenous; MTX=methotrexate; SC=subcutaneous; VdH=van der Heijde.

combinations of treatments with methotrexate (azathioprine, ciclosporin, tocilizumab 8 mg/kg) had a statistically significant increase in the rate of withdrawals due to adverse events compared with oral methotrexate in either the methotrexate naive or methotrexate inadequate response populations.

Completeness and applicability of evidence

We did not evaluate the effect of glucocorticoids, which are known to have a disease modifying effect, particularly in early rheumatoid arthritis.[32] We did not exclude trials with corticosteroids, however, so our findings relate to the effects of DMARD therapy independent of a corticosteroid effect. Our results should not be generalized to patients who have had an inadequate response to biologic therapy, as we did not include these trials. Our study was also not designed to directly compare treatment strategies. Specifically, we did not directly compare the approach of starting with methotrexate monotherapy in methotrexate naive patients and progressing to triple therapy versus the strategy of starting with triple therapy directly. The estimates of absolute risk with each treatment can help to inform this decision. On the basis of the included trials, about 40% of patients naive to methotrexate are expected to have an ACR50 response to oral methotrexate, compared with 60% for triple therapy. Patients may accept this difference in risk and choose methotrexate monotherapy as initial treatment, reserving combination therapy for if they fail to respond adequately. Triple therapy was also associated with an increase in total gastrointestinal events in methotrexate naive patients, which may influence patients' decisions.

Strengths and weaknesses of review

This review included 158 trials with more than 37 000 patients. We used a rigorous approach to identification of trials and abstraction of outcomes, so we have confidence that the results encompass the best evidence from randomized controlled trials of the comparative benefits and harms of the treatments of interest.

The extent to which indirect evidence is considered can affect the results of a network meta-analysis.[33] With our search strategy, we included all direct and first order indirect evidence comparing the treatments of interest; we did not attempt to capture all second order indirect evidence. The contribution of the indirect evidence to the overall estimate from the network meta-analysis decreases quite rapidly as the "order" of the comparison increases.[33] In addition, most of the treatments that would potentially form second order indirect evidence were monotherapy with conventional synthetic DMARDs, for which the few trials that exist are generally small and have rarely measured ACR responses.[32] We therefore expect a minimal effect of the exclusion of these trials.

An "early escape" design was common in trials of methotrexate plus biologic DMARDs and methotrexate plus tofacitinib, particularly in more recent trials. Although this allows trials to ethically include a placebo arm, it presents challenges in interpreting and synthesizing the results. The proportion of patients remaining on the control treatment at the end of the trial can be very low. We chose to extract end of trial data for the efficacy outcomes, considering patients who required rescue treatment to be treatment failures. We included a sensitivity analysis for ACR50 response using pre-rescue data and found few differences in treatment effects. Synthesizing adverse events in early escape trials is also challenging. Patients who cross over from placebo to active therapy often represent substantial patient years of exposure; excluding these patients may obscure important safety signals. We therefore choose to summarize all toxicity data as rate ratios with exposure adjusted estimates, using the on-treatment data from early escape trials. This could potentially bias the estimates, as patients who cross over may differ in certain ways from patients assigned to the original treatment. We believed this method to have less potential bias than excluding the patients who crossed over and excluding trials that reported only exposure adjusted data.

Through meta-regression, we showed that treatments in trials of patients with an inadequate response to methotrexate were associated with odds ratios for ACR50 response that were twice as high as those in trials of methotrexate naive patients. Thus, previous response to methotrexate is a strong effect modifier of the clinical response, and pooling studies in methotrexate naive and methotrexate inadequate response patients will yield biased estimates that are difficult to relate to clinical practice. This supports our decision to analyze trials of methotrexate naive and methotrexate inadequate response patients separately. We used end of trial data for all outcomes, pooling studies with variable follow-up (from three months to two years). For ACR50 response, our results were robust to sensitivity analyses using six or 12 month data, and we did not detect an association between trial duration and treatment effects through meta-regression. We had too few trials available for radiographic progression to be able to do meaningful sensitivity analyses or meta-regression, so this may represent a source of heterogeneity.

Agreements and disagreements with other studies or reviews

Multiple network meta-analyses of biologic therapy in rheumatoid arthritis, including a Cochrane review, have been reported.[4 34] To our knowledge, this is the first review that has systematically compared all methotrexate based conventional synthetic DMARD and biologic DMARD/tofacitinib treatment approaches. A previous network meta-analysis by Graudal and colleagues compared combination DMARD therapy with conventional synthetic DMARDs and biologic DMARDs for radiographic outcomes.[35] Overall, Graudal found that one conventional synthetic DMARD plus one biologic DMARD was not superior to combination therapy with two or three conventional synthetic DMARDs for radiographic outcomes. Several important differences from our study exist. Firstly, we evaluated a range of outcomes beyond radiographic progression, covering

multiple domains relevant to decision making. Secondly, Graudal grouped conventional synthetic DMARD combination therapy according to the number of drugs (two or three), whereas each biologic agent was considered separately. Grouping DMARD combinations that are commonly used with those that are rarely used (for example, combinations with bucillamine or auranofin) adds heterogeneity to the estimates and makes application of the results to clinical practice difficult. Trials in DMARD naive and DMARD inadequate response patients were also grouped, which, as we showed for ACR50 response, may bias the estimated treatment effects. Finally, Graudal included trials in which corticosteroids were part of the intervention (that is, applied differently between arms). The results therefore address a different research question.

Other traditional (non-network) systematic reviews have evaluated combination therapy with conventional synthetic DMARDs.[5 36 37] The reviews differed in the outcomes considered and inclusion criteria, particularly around the inclusion and exclusion of interventions with corticosteroids. Combined withdrawal due to inefficacy or adverse events was used as the primary outcome for two of the reviews,[5 37] as it is commonly reported. Trials are not designed around this outcome, however, and it does not allow the separation of benefits and harms necessary to inform clinical decisions.

In contrast to other systematic reviews, our review evaluated the risk of bias separately for different outcomes, which is the approach recommended by GRADE.[38] We also used recently published GRADE guidance for grading the quality of the evidence.[21] Although this approach requires subjective decisions, it should increase the transparency of these choices.

Implications for practice
On the basis of all available direct and indirect evidence, our results suggest that triple therapy (methotrexate plus sulfasalazine plus hydroxychloroquine) is effective in both methotrexate naive and methotrexate inadequate response patients and not statistically different from methotrexate plus biologic therapy for controlling disease activity. As triple therapy costs 10-20 times less than biologic therapy and is not currently recommended strongly by international guidelines, this has important policy implications. Specifically, our results suggest that triple therapy should be considered as a low cost, effective treatment option either as initial treatment or after an inadequate response to methotrexate.

AUTHOR AFFILIATIONS
[1]Department of Medicine, University of Calgary, Calgary, AB, Canada, T2N4Z6

[2]McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, AB, Canada, T2N4Z6

[3]Institute of Health, Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada, M5T3M6

[4]Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada, T2N4Z6

[5]Department of Medicine and Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada, M5G2C4

[6]Toronto General Research Institute, University Health Network, Toronto, ON, Canada, M6J3S3

[7]Mount Sinai Hospital, Division of Rheumatology, Toronto, ON, Canada, M5T3L9

1   Smolen JS, Landewé R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73:492-509. doi:10.1136/annrheumdis-2013-204573.
2   Singh JA, Saag KG, Bridges SL Jr, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheum* 2016;68:1-26.
3   Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011;14:417-28. doi:10.1016/j.jval.2011.04.002.
4   Singh JA, Christensen R, Wells GA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. *Cochrane Database Syst Rev* 2009;4:CD007848.
5   Katchamart W, Trudeau J, Phumethum V, Bombardier C. Methotrexate monotherapy versus methotrexate combination therapy with non-biologic disease modifying anti-rheumatic drugs for rheumatoid arthritis. *Cochrane Database Syst Rev* 2010;4:CD008495.

6   O'Dell JR, Mikuls TR, Taylor TH, et al. CSP 551 RACAT Investigators. Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med* 2013;369:307-18. doi:10.1056/NEJMoa1303006.

7   van Vollenhoven RF, Geborek P, Forslind K, et al. Swefot study group. Conventional combination treatment versus biological treatment in methotrexate-refractory early rheumatoid arthritis: 2 year follow-up of the randomised, non-blinded, parallel-group Swefot trial. *Lancet* 2012;379:1712-20. doi:10.1016/S0140-6736(12)60027-0.

8   Moreland LW, O'Dell JR, Paulus HE, et al. TEAR Investigators. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis Rheum* 2012;64:2824-35. doi:10.1002/art.34498.

9   Hazlewood G, Barnabe C, Tomlinson G, Marshall D, Bombardier C. Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic disease modifying anti-rheumatic drugs for rheumatoid arthritis: A network meta-analysis (Protocol). *Cochrane Database Syst Rev* 2013;1:CD010227.

10   Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580-8. doi:10.1136/ard.2010.138461.

11   Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24. doi:10.1002/art.1780310302.

12   Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 Revision of diagnostic criteria for rheumatoid arthritis. *Bull Rheum Dis* 1958;9:175-6.

13   Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727-35. doi:10.1002/art.1780380602.

14   Ory PA. Interpreting radiographic data in rheumatoid arthritis. *Ann Rheum Dis* 2003;62:597-604. doi:10.1136/ard.62.7.597.

15   Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928.

16   Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607-17. doi:10.1177/0272989X12458724.

17   Ades AE, Sculpher M, Sutton A, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:1-19. doi:10.2165/00019053-200624010-00001.

18   Jonas DE, Wilkins TM, Bangdiwala S, et al. *Findings of Bayesian Mixed Treatment Comparison Meta-Analyses: Comparison and Exploration Using Real-World Trial Data and Simulation.*Agency for Healthcare Research and Quality, 2013.

19   Brooks S, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *J Comput Graph Stat* 1998;7:434-55.

20   Plummer M. rjags: Bayesian Graphical Models using MCMC. R package version 3-14. 2014. https://libraries.io/cran/rjags/3-14.

21   Puhan MA, Schünemann HJ, Murad MH, et al. GRADE Working Group. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;349:g5630. doi:10.1136/bmj.g5630.

22   Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44. doi:10.1002/sim.3767.

23   Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Stat Med* 2005;24:2401-28. doi:10.1002/sim.2112.

24   Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol* 2012;41:818-27. doi:10.1093/ije/dys041.

25   Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *J Clin Epidemiol* 2015;68:52-60. doi:10.1016/j.jclinepi.2014.08.012.

26   van Vollenhoven RF, Fleischmann R, Cohen S, et al. ORAL Standard Investigators. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. *N Engl J Med* 2012;367:508-19. doi:10.1056/NEJMoa1112072.

27   Schiff M, Weinblatt ME, Valente R, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. *Ann Rheum Dis* 2014;73:86-94. doi:10.1136/annrheumdis-2013-203843.

28   Schiff M, Keiserman M, Codding C, et al. Efficacy and safety of abatacept or infliximab vs placebo in ATTEST: a phase III, multi-centre, randomised, double-blind, placebo-controlled study in patients with rheumatoid arthritis and an inadequate response to methotrexate. *Ann Rheum Dis* 2008;67:1096-103. doi:10.1136/ard.2007.080002.

29   Cuomo G, Molinaro G, La Montagna G, Migliaresi S, Valentini G. [A comparison between the Simplified Disease Activity Index (SDAI) and the Disease Activity Score (DAS28) as measure of response to treatment in patients undergoing different therapeutic regimens][Italian]. *Reumatismo* 2006;58:22-5.

30   Joo K, Heejung K, Lim MJ, Kwon SR, Park W. Safety and efficacy of TNFa inhibitor versus leflunomide in patients with rheumatoid arthritis inadequately responding to methotrexate in Korea. *Int J Rheum Dis* 2012;15:53-4.

31   Bruynesteyn K, van der Heijde D, Boers M, et al. Determination of the minimal clinically important difference in rheumatoid arthritis joint damage of the Sharp/van der Heijde and Larsen/Scott scoring methods by clinical experts and comparison with the smallest detectable difference. *Arthritis Rheum* 2002;46:913-20. doi:10.1002/art.10190.

32   Gaujoux-Viala C, Nam J, Ramiro S, et al. Efficacy of conventional synthetic disease-modifying antirheumatic drugs, glucocorticoids and tofacitinib: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. *Ann Rheum Dis* 2014;73:510-5. doi:10.1136/annrheumdis-2013-204588.

33   Hawkins N, Scott DA, Woods B. How far do you go? Efficient searching for indirect evidence. *Med Decis Making* 2009;29:273-81. doi:10.1177/0272989X08330120.

34   Thorlund K, Druyts E, Aviña-Zubieta JA, Wu P, Mills EJ. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: an overview of recurrent methodological shortcomings. *Ann Rheum Dis* 2013;72:1524-35. doi:10.1136/annrheumdis-2012-201574.

35   Graudal N, Hubeck-Graudal T, Tarp S, Christensen R, Jürgens G. Effect of combination therapy on joint destruction in rheumatoid arthritis: a network meta-analysis of randomized controlled trials. *PLoS One* 2014;9:e106408. doi:10.1371/journal.pone.0106408.

36   Graudal N, Jürgens G. Similar effects of disease-modifying antirheumatic drugs, glucocorticoids, and biologic agents on radiographic progression in rheumatoid arthritis: meta-analysis of 70 randomized placebo-controlled or drug-controlled studies, including 112 comparisons. *Arthritis Rheum* 2010;62:2852-63. doi:10.1002/art.27592.

37   Choy EH, Smith C, Doré CJ, Scott DL. A meta-analysis of the efficacy and toxicity of combining disease-modifying anti-rheumatic drugs in rheumatoid arthritis based on patient withdrawal. *Rheumatology (Oxford)* 2005;44:1414-21. doi:10.1093/rheumatology/kei031.

38   Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol* 2011;64:407-15. doi:10.1016/j.jclinepi.2010.07.017.

Web Extra

Appendix A

Appendix B

Appendix C

Appendix D

Appendix E