Drug Therapy for Early Rheumatoid Arthritis: A Systematic Review Update
Comparative Effectiveness Review
Number 211

Drug Therapy for Early Rheumatoid Arthritis: 
A Systematic Review Update

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

and

Patient-Centered Outcomes Research Institute
1828 L Street, NW, Ste. 900
Washington, DC 20036
www.pcori.org

Contract No. 290-2015-00011-I

Prepared by:
RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center
Research Triangle Park, NC

Investigators:
Katrina E. Donahue, M.D., M.P.H.
Gerald Gartlehner, M.D., M.P.H.
Elizabeth R. Schulman, M.D.
Beth Jonas, M.D., FACR
Emmanuel Coker-Schwimmer, M.P.H.
Sheila V. Patel, B.S.P.H.
Rachel Palmieri Weber, Ph.D.
Kathleen N. Lohr, Ph.D., M.Phil., M.A.
Carla Bann, Ph.D.
Meera Viswanathan, Ph.D.

AHRQ Publication No. 18-EHC015-EF
PCORI Publication No. 2018-SR-02
July 2018
Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States.

The Patient-Centered Research Outcomes Institute (PCORI) was established to fund research that can help patients and those who care for them make better informed decisions about the health care choices they face every day. PCORI partnered with AHRQ to help fulfill PCORI’s authorizing mandate to engage in evidence synthesis and make information from comparative effectiveness research more available to patients and providers. PCORI identifies topics for review based on broad stakeholder interest. After identifying specific topics, multistakeholder virtual workshops are held by PCORI to inform the individual research protocols.

The reports and assessments provide organizations, patients, clinicians, and caregivers with comprehensive, evidence-based information on common medical conditions and new health care technologies and strategies. They also identify research gaps in the selected scientific area, identify methodological and scientific weaknesses, suggest research needs, and move the field forward through an unbiased, evidence-based assessment of the available literature. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ, and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review and public comment prior to their release as a final report.

AHRQ expects that the EPC evidence reports and technology assessments, when appropriate, will inform patients and caregivers, individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

If you have comments on this evidence report, they may be sent by mail to the Task Order Officer: Aysegul Gozu, M.D., M.P.H., Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Gopal Khanna M.B.A.
Director
Agency for Healthcare Research and Quality

Arlene S. Bierman, M.D., M.S.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Joe V. Selby, M.D., M.P.H.
Executive Director
PCORI

Diane E. Bild, M.D., M.P.H.
Acting Chief Science Officer
PCORI
Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project and deeply appreciate their considerable support, commitment, and contributions: Aysegul Gozu, M.D., M.P.H., our AHRQ Task Order Officer; Jennifer Croswell, M.D., M.P.H., our PCORI Senior Program Officer; Jennie Bowen, M.P.H., our PCORI Program Officer; John Williams, M.D., M.H.S., our Associate Editor; Loraine Monroe, our publications specialist; Sharon Barrell, M.A., our editor; Carol Woodell, B.S.P.H., our project manager; Charli Randolph, B.A., who assisted with data abstraction, data accuracy checking, and full-text article retrieval; Claire Baker, who assisted with data accuracy checking and full-text article retrieval; Cassandra J. Bamhart, M.P.H., who assisted with full-text article retrieval; Ursula Griebler, Ph.D., M.P.H., who helped assign risk of bias ratings to our included studies; Christina Kien, M.A., who helped assign risk of bias ratings to our included studies; Gernot Wagner, M.D., who provided instruction to Mr. Coker-Schwimmer on using data visualization software to read figure-only study data and reviewed data retrieved this way; and Laurie Leadbetter, M.S.L.S., and Christiane Voisin, M.S.L.S., who provided library services.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence Summary</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Condition</td>
<td>1</td>
</tr>
<tr>
<td>Etiology</td>
<td>1</td>
</tr>
<tr>
<td>Burden of Disease</td>
<td>1</td>
</tr>
<tr>
<td>Definitions of Early RA and Challenges With the Definitions</td>
<td>1</td>
</tr>
<tr>
<td>Current Practice and Treatment Strategies</td>
<td>2</td>
</tr>
<tr>
<td>Scope and Key Questions</td>
<td>4</td>
</tr>
<tr>
<td>Scope of the Review</td>
<td>4</td>
</tr>
<tr>
<td>Key Questions</td>
<td>5</td>
</tr>
<tr>
<td>Contextual Questions</td>
<td>5</td>
</tr>
<tr>
<td>Analytic Framework</td>
<td>5</td>
</tr>
<tr>
<td>Organization of This Report</td>
<td>6</td>
</tr>
<tr>
<td>Methods</td>
<td>7</td>
</tr>
<tr>
<td>Criteria for Inclusion/Exclusion of Studies in the Review</td>
<td>7</td>
</tr>
<tr>
<td>Searching for the Evidence: Literature Search Strategies</td>
<td>8</td>
</tr>
<tr>
<td>To Answer the Key Questions</td>
<td>8</td>
</tr>
<tr>
<td>Literature Review, Data Abstraction, and Data Management</td>
<td>9</td>
</tr>
<tr>
<td>Assessment of Methodological Risk of Bias of Individual Studies</td>
<td>11</td>
</tr>
<tr>
<td>Data Synthesis</td>
<td>12</td>
</tr>
<tr>
<td>Grading the Strength of Evidence for Major Comparisons and Outcomes</td>
<td>13</td>
</tr>
<tr>
<td>Assessing Applicability</td>
<td>14</td>
</tr>
<tr>
<td>Peer Review and Public Commentary</td>
<td>14</td>
</tr>
<tr>
<td>Results</td>
<td>15</td>
</tr>
<tr>
<td>Organization of the Results</td>
<td>15</td>
</tr>
<tr>
<td>Search Results</td>
<td>15</td>
</tr>
<tr>
<td>Characteristics of Included Studies</td>
<td>15</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>21</td>
</tr>
<tr>
<td>csDMARD Studies</td>
<td>22</td>
</tr>
<tr>
<td>KQ 1: Comparative Benefits of Drug Therapies for Patients With Early RA in Relation to Disease Activity, Progression of Radiographic Joint Damage, or Remission</td>
<td>27</td>
</tr>
<tr>
<td>Key Points</td>
<td>27</td>
</tr>
<tr>
<td>Detailed Synthesis</td>
<td>28</td>
</tr>
<tr>
<td>KQ 2: Comparative Benefits of Drug Therapies for Patients With Early RA in Relation to Patient-Reported Symptoms, Functional Capacity, or Quality of Life</td>
<td>59</td>
</tr>
<tr>
<td>Key Points</td>
<td>60</td>
</tr>
<tr>
<td>Detailed Synthesis</td>
<td>74</td>
</tr>
<tr>
<td>KQ 3: Comparative Harms of Drug Therapies for Patients With Early RA in Relation to Harms, Tolerability, Patient Adherence, or Adverse Effects</td>
<td>82</td>
</tr>
<tr>
<td>Key Points</td>
<td>83</td>
</tr>
<tr>
<td>Detailed Synthesis</td>
<td>83</td>
</tr>
<tr>
<td>KQ 4: Comparative Benefits and Harms in Subgroups of Patients</td>
<td>114</td>
</tr>
<tr>
<td>Key Points</td>
<td>114</td>
</tr>
<tr>
<td>Detailed Synthesis</td>
<td>115</td>
</tr>
</tbody>
</table>
Discussion................................................................................................................................... 117
Overview of Key Findings...........................................................................................................117
Findings in Relationship to What Is Already Known...................................................................126
Applicability ................................................................................................................................127
Contextual Questions...................................................................................................................128
  Contextual Question 1: Does treatment of early RA improve disease trajectory and
disease outcomes compared with the trajectory or outcomes of treatment of established
RA? ..................................................................................................................................128
  Contextual Question 2: What barriers prevent individuals with early RA from obtaining
access to indicated drug therapies? ..................................................................................128
Limitations ..................................................................................................................................129
Research Needs ..........................................................................................................................130
Conclusions ..............................................................................................................................130
Abbreviations and Acronyms .................................................................................................... 131
Report References .................................................................................................................. Ref-1

Tables
Table A. Summary of characteristics of this systematic review on treatment of patients
with early rheumatoid arthritis.................................................................................. ES-1
Table B. Benefits and harms of csDMARDs versus csDMARD or tsDMARD for treatment of
patients with early rheumatoid arthritis .............................................................. ES-6
Table C. Benefits and harms of biologic DMARDs for early RA treatment ............................ ES-8
Table 1. Corticosteroids and disease-modifying antirheumatic drugs (DMARDs) approved
by the U.S. Food and Drug Administration ............................................................... 3
Table 2. Eligibility criteria for review of treatments for early rheumatoid arthritis ................. 7
Table 3. Outcomes and hierarchy of preferred measures for data abstraction ......................... 10
Table 4. Definitions of the grades of overall strength of evidence ............................................. 13
Table 5. Number of studies included for each KQ, by drug therapy group, comparison type,
and study design............................................................................................................ 18
Table 6. Characteristics of included studies ............................................................................. 19
Table 7. Disease activity, response, and radiographic progression ........................................... 29
Table 8. Studies included in KQ 1 network meta-analyses ........................................................ 42
Table 9. Results for patient-reported outcomes, functional status, and quality of life .......... 61
Table 10. Discontinuation rates and adverse events ................................................................. 84
Table 11. Studies included in KQ 3 network meta-analysis ........................................................ 98
Table 12. Summary of findings about benefits and harms of treatments for early rheumatoid
arthritis with strength of evidence grades ................................................................... 117

Figures
Figure 1. Analytic framework for drug therapy for early rheumatoid arthritis ....................... 6
Figure 2. Summary of literature search flow and yield for early rheumatoid arthritis .......... 16
Figure 3. Network diagram for network meta-analysis: ACR50 response rates .................. 41
Figure 4. Network diagram for network meta-analysis: change from baseline in
radiographic joint damage score ............................................................................... 41
Figure 5. Forest plot for network meta-analysis of ACR50 response rates: MTX plus TNF biologic versus TNF biologic ................................................................. 45
Figure 6. Forest plot for network meta-analysis of change from baseline in radiographic joint damage score: MTX plus TNF biologic versus TNF biologic ........................................ 46
Figure 7. Forest plot for network meta-analysis of ACR50 response rates: MTX plus non-TNF versus non-TNF biologic ................................................................. 47
Figure 8. Forest plot for network meta-analysis of change from baseline in radiographic joint damage score: MTX plus non-TNF versus non-TNF biologic .................. 47
Figure 9. Forest plot for network meta-analysis of ACR50 response rates: Comparison of TNF plus MTX with MTX only .................................................................. 49
Figure 10. Forest plot for network meta-analysis of change from baseline in radiographic joint damage score: Comparison of TNF combined therapies with MTX only ........ 49
Figure 11. Forest plot for network meta-analysis of ACR50 response rates: Non-TNF biologic plus MTX versus MTX ................................................................. 53
Figure 12. Forest plot for change from baseline in radiographic joint damage score: Non-TNF biologic plus MTX versus MTX ..................................................... 53
Figure 13. Forest plot for network meta-analysis of ACR50 response rates: TNF biologic versus non-TNF biologic ................................................................. 55
Figure 14. Forest plot for change from baseline in radiographic joint damage score: TNF biologic versus non-TNF biologic ..................................................... 55
Figure 15. Forest plot for network meta-analysis of ACR50 response rates: TNF biologic versus TNF biologic ................................................................. 56
Figure 16. Forest plot for change from baseline in radiographic joint damage score: TNF biologic versus TNF biologic ............................................................. 57
Figure 17. Forest plot for network meta-analysis of ACR50 response rates: Non-TNF biologic versus non-TNF biologic ............................................................. 57
Figure 18. Forest plot for change from baseline in radiographic joint damage score: Non-TNF biologic versus non-TNF biologic ..................................................... 58
Figure 19. Network diagram for network meta-analysis: All discontinuations and discontinuations due to adverse events ................................................................. 97
Figure 20. Forest plot for network meta-analysis: All discontinuations, csDMARD monotherapy versus csDMARD monotherapy ........................................... 100
Figure 21. Forest plot for network meta-analysis: Discontinuations due to adverse events, csDMARD monotherapy versus csDMARD monotherapy ............... 100
Figure 22. Forest plot for network meta-analysis: All discontinuations, TNF + MTX versus TNF ................................................................. 101
Figure 23. Forest plot for network meta-analysis: Discontinuations due to adverse events, TNF + MTX versus TNF ................................................................. 101
Figure 24. Forest plot for network meta-analysis: All discontinuations, non-TNF + MTX versus non-TNF ................................................................. 102
Figure 25. Forest plot for network meta-analysis: Discontinuations due to adverse events, non-TNF + MTX versus non-TNF ................................................................. 103
Figure 26. Forest plot for network meta-analysis: All discontinuations, TNF + MTX versus csDMARD ................................................................. 104
Figure 27. Forest plot for network meta-analysis: Discontinuations due to adverse events, TNF + MTX versus csDMARD ................................................................. 105
Figure 28. Forest plot for network meta-analysis: All discontinuations, non-TNF + MTX versus csDMARD ................................................................. 107
Figure 29. Forest plot for network meta-analysis: Discontinuations due to adverse events, non-TNF + MTX versus csDMARD ................................................................. 108
Figure 30. Forest plot for network meta-analysis: All discontinuations, TNF versus non-TNF .............................................................................. 109
Figure 31. Forest plot for network meta-analysis: Discontinuations due to adverse events, TNF versus non-TNF .............................................................................. 110
Figure 32. Forest plot for network meta-analysis: All discontinuations, TNF versus TNF ................................................................. 111
Figure 33. Forest plot for network meta-analysis: Discontinuations due to adverse events, TNF versus TNF .............................................................................. 112
Figure 34. Forest plot for network meta-analysis: All discontinuations, non-TNF versus non-TNF ................................................................. 113
Figure 35. Forest plot for network meta-analysis: Discontinuations due to adverse events, non-TNF versus non-TNF .............................................................................. 113

Appendixes
Appendix A. Search Strings
Appendix B. Excluded Articles
Appendix C. Detailed Evidence Table
Appendix D. Risk of Bias Ratings and Rationales for Included Studies
Appendix E. Strength of Evidence for Key Questions 1-4 Outcomes
Appendix F. Eligible Clinical and Self-Reported Scales and Instruments Commonly Used in Eligible Studies of Drug Therapy for Rheumatoid Arthritis
Appendix G. Tests of Consistency for Main Network Meta-Analyses
Appendix H. Supplementary Primary Network Meta-Analyses
Appendix I. Sensitivity Analyses for Network Meta-Analyses
Appendix J. Expert Guidance and Review
Appendix K. PCORI Methodology Standards Checklist SER Update
Evidence Summary

Introduction and Methods

This systematic review updates a 2012 report that evaluated the benefits and harms of drug therapies for adults with rheumatoid arthritis (RA).\(^1\) This updated review, however, has a targeted scope focusing solely on patients with early RA. Early RA has no formal consensus definition. Based on guidance from a recent task force of experts,\(^2\) we define early RA as no more than 1 year of diagnosed disease. Our findings should be considered applicable only to patients with early RA.

The U.S. Food and Drug Administration (FDA) has approved several drug therapy groups for treating patients with RA. Corticosteroids and conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) have been prescribed the longest. Targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) were approved more recently. Additionally, many trials or observational studies in this review evaluated mainly FDA-approved biologic drugs (both tumor necrosis factor [TNF] and non-TNF drugs). The FDA has approved numerous biosimilars.

We evaluated the benefits and harms of multiple drug monotherapies, combination therapies, and different treatment strategies to determine whether therapeutic approaches differ in their ability to affect important outcomes for patients with early RA. The benefits and harms included (1) reduced disease activity, decreased progression of joint damage, or remission; (2) improved functional capacity or quality of life; (3) harms such as tolerability, serious adverse events, and adverse effects; and (4) benefits and harms among patient subgroups (based on disease activity, prior RA therapy, demographics, or presence of other diseases with or without treatment). Additional details about this systematic review are described in Table A. Two Contextual Questions were also examined: (1) Does treatment of early RA improve disease trajectory and disease outcomes compared with the trajectory or outcomes of treatment of established RA? And (2) What barriers prevent individuals with early RA from obtaining access to indicated drug therapies?

We synthesized the literature qualitatively within and between corticosteroids and classes of disease-modifying antirheumatic drugs (DMARDs), including csDMARDs, tsDMARDs, TNF and non-TNF biologics, and biosimilars. Additionally, combination treatment strategies were examined. We conducted network meta-analysis for five outcomes: American College of Rheumatology 50 percent improvement (ACR50), remission based on Disease Activity Score (DAS), radiographic joint damage, all discontinuations, and discontinuations due to adverse events.

Table A. Summary of characteristics of this systematic review on treatment of patients with early rheumatoid arthritis

<table>
<thead>
<tr>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Inclusion Criteria:</strong> Adult outpatients, 19 years of age or older, with an early RA diagnosis, defined as 1 year or less from disease diagnosis</td>
</tr>
<tr>
<td><strong>Key Exclusion Criteria:</strong> Adolescents and adults with RA greater than 1 year from diagnosis</td>
</tr>
</tbody>
</table>

Note: The references for the Evidence Summary are included in the reference list that follows the appendixes.
Drug Therapies Approved by FDA for RA Included in the Review

Corticosteroid: Methylprednisolone, prednisone (PRED), prednisolone (PNL)

Conventional synthetic DMARD (csDMARD): Hydroxychloroquine (HCQ), leflunomide (LEF), methotrexate (MTX), sulfasalazine (SSZ)

Tumor necrosis factor (TNF) biologic DMARD: Adalimumab (ADA), certolizumab pegol (CZP), etanercept (ETN), golimumab (GOL), infliximab (IFX)

Non-TNF biologic DMARD: Abatacept (ABA), rituximab (RIT), sarilumab (SAR), tocilizumab (TCZ)

Targeted synthetic DMARD (tsDMARD): Tofacitinib (TOF)

Biosimilars: ADA-atto, IFX-dyyb, IFX-abda, ETN-szxs

Key Questions Covered by the Review

1. Benefits of drug therapies including reducing disease activity, slowing or limiting the progression of joint damage, or inducing remission

   Clinical tools including:
   - ACR 20/50/70
   - DAS
   - Sharp Score

2. Benefits of drug therapies including improving patient-reported symptoms, functional capacity, or quality of life

   Clinical tools including:
   - HAQ
   - SF-36

3. Harms of drug therapies including tolerability, patient adherence, and adverse effects

   Harms including:
   - Overall discontinuations
   - Discontinuations attributable to AEs
   - Serious AEs
   - Specific AEs

4. Benefits and harms of drug therapies in subgroups of patients

   Subgroups of patients defined by:
   - Age
   - Sex
   - Race or ethnicity
   - Disease activity
   - Prior treatment
   - Concomitant therapies
   - Coexisting conditions

Timing of Review

Beginning Search Date: January 2011

End Search Date: October 5, 2017
Overview of Important Studies Underway

<table>
<thead>
<tr>
<th>Study Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Six trials either ongoing or completed, but findings not yet published.</td>
</tr>
<tr>
<td>• One trial of ETN plus MTX versus a treat-to-target strategy with initial</td>
</tr>
<tr>
<td>MTX and later combination DMARD treatment</td>
</tr>
<tr>
<td>• One trial of TCZ plus MTX versus TCZ</td>
</tr>
<tr>
<td>• One trial of TCZ plus MTX versus MTX</td>
</tr>
<tr>
<td>• One trial of ETN plus MTX versus MTX</td>
</tr>
<tr>
<td>• One trial of csDMARD combination therapy versus csDMARD monotherapy</td>
</tr>
<tr>
<td>• One single-arm study of golimumab.</td>
</tr>
</tbody>
</table>

a New medications that FDA has approved since the prior report
b Sharp-van der Heijde method for scoring radiographs
c As defined by FDA: Life-threatening, requires hospitalization, leads to lasting disability or congenital anomaly, or jeopardizes the patient in any serious way
d Rash, upper respiratory tract infection, nausea, pruritus, headache, diarrhea, dizziness, abdominal pain, bronchitis, leukopenia, injection site reactions

ACR 20/50/70 = American College of Rheumatology 20/50/70% improvement from baseline; AE = adverse event; csDMARD = conventional synthetic disease-modifying antirheumatic drug; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; FDA = U.S. Food and Drug Administration; HAQ = Health Assessment Questionnaire; MTX = methotrexate; RA = rheumatoid arthritis; SF-36 = Medical Outcomes Study Short Form 36 Health Survey; TCZ = tocilizumab.

Results and Key Findings

We included 49 studies (reported in 124 published articles) that provided data on at least one of the review’s Key Questions. Of these studies, 41 were randomized controlled trials (RCTs), 4 were observational studies with control groups, and 4 were single-arm observational studies that we included only for evaluating harms of treatment.

We rated a majority of these studies as low or medium risk of bias. We rated 16 studies as high risk of bias for at least some of the eligible outcomes they reported. Studies (n=2) or study outcomes rated high risk of bias were excluded from analyses and used only in sensitivity analyses for the network meta-analysis. We graded strength of evidence (SOE) for numerous outcomes in studies for these drug classes and therapeutic approaches (except that the single-arm observational studies were not included in the SOE assessments).

The range of mean (or median) disease durations across all 49 included studies was 2 weeks to 12 months. Prior treatment use varied widely across drug therapy categories. Among all 49 included studies, five studies did not report any details about prior treatment use, leaving 44 studies that did. Of these, 36 enrolled methotrexate (MTX)-naïve patient samples, and the remaining eight studies enrolled patients with at least some prior csDMARD use (including MTX).

In four of these eight studies, prior use of any csDMARDs ranged from 13 to 48 percent. The other four enrolled samples that were entirely csDMARD resistant. Among the 15 studies analyzed in our primary or sensitivity network meta-analysis (NWMA), five enrolled patients with some prior csDMARD use other than MTX, and three did not report whether patients had used other csDMARDs.

Five of the eight studies enrolled samples that had previously used MTX specifically: 5819 and 7920 percent of patients in two studies, and three studies (all trials) enrolling samples that were entirely MTX resistant (i.e., 100% prior use).5-10

All included studies enrolled patients with moderate to high disease activity at baseline as measured with mean or median Disease Activity Score (DAS) 28 scores (range of 0 to 10);
DAS28 scores in these studies ranged from 3.4 to 7.1. A DAS28 score of 3.2 is the threshold for low disease activity; a score exceeding 5.1 translates to high disease activity. Additional detailed information about the DAS28 is available in Appendix F.

More than one-half (ranging from 53% to 83%) of the patient population was women. The mean age range was 46 to 64 years. Study durations ranged from 6 months to 15 years.

We grouped studies based on the primary drug therapy of interest, ordered from oldest (corticosteroids and csDMARDs) to newest (TNF or non-TNF biologics), and then the most complex (combination therapies). We describe the main findings for each group below.

**Corticosteroids:** Eight RCTs evaluated corticosteroids, and one single-arm observational study provided additional data on harms. A corticosteroid, when taken with a csDMARD (usually MTX), led to higher remission rates than the csDMARD alone (from 44.8% to 76.7% for combination therapy and 27.8% to 33.3% for MTX monotherapy) (low SOE). Groups did not differ significantly in terms of serious adverse events and discontinuations attributable to adverse events (graded moderate and low SOE, respectively). We could not draw conclusions about disease activity, radiographic changes, or functional capacity because evidence was insufficient.

**csDMARDs and tsDMARDs:** Twelve RCTs, two observational studies with control groups, and four single-arm observational studies evaluated csDMARDs; only one of these studies compared a tsDMARD (tofacitinib) with a csDMARD (MTX) (Table B). These studies predominantly compared sulfasalazine plus MTX with MTX only. When comparing various csDMARD combination therapies with csDMARD monotherapies, we concluded that groups did not differ in response, remission, functional capacity, serious adverse events, or discontinuations attributable to adverse events (low SOE). Evidence was insufficient to draw conclusions about radiographic changes, csDMARD monotherapies compared with other csDMARD monotherapies, or tofacitinib compared with MTX.

When comparing csDMARD (MTX) plus TNF biologic therapy (adalimumab [ADA]) with ADA only, we concluded that combination treatment led to greater response and remission, less radiographic progression, and greater improvement in functional capacity (moderate SOE) (Table C). The groups also did not differ in serious adverse events or discontinuations attributable to adverse events (moderate SOE).

Treatment with a csDMARD (MTX) plus a non-TNF biologic (TCZ) led to greater remission than TCZ biologic monotherapy, respectively (low SOE), but groups receiving treatment with MTX plus another non-TNF biologic (abatacept [ABA]) did not differ in response or remission from those receiving ABA monotherapy (Table C). Groups receiving MTX plus ABA did not differ from those receiving ABA monotherapy in functional activity (low SOE). The groups also did not differ in serious adverse events or discontinuations attributable to adverse events (low SOE for ABA and moderate SOE for TCZ). Evidence was insufficient to draw conclusions about disease activity for these comparisons or about functional capacity for MTX plus TCZ compared with TCZ.

**Biologic DMARDs:** Twenty-two RCTs and one single-arm observational study evaluated TNF and non-TNF biologic drugs for treating patients with early RA. Of these, 22 evaluated disease activity, functional capacity, and harms outcomes (Table C). The combination of either a TNF or a non-TNF biologic with MTX, when compared with MTX alone, generally reduced disease activity (mostly moderate and low SOE) and led to higher rates of remission (all moderate and low SOE) and less radiographic progression (mostly moderate and low SOE). Network meta-analyses and head-to-head trials found higher ACR50 response for combination therapy of biologic DMARDs plus MTX than MTX monotherapy (NWMA range of relative
risks [RRs], 1.20 [95% confidence interval [CI], 1.04 to 1.38] to 1.57 [95% CI, 1.30 to 1.88]). The groups did not differ with respect to harms (all low SOE).
<table>
<thead>
<tr>
<th>Outcome Type</th>
<th>Specific Outcome</th>
<th>Results</th>
<th>Strength of Evidence</th>
<th>Summary of Rationale for Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>No significant difference between SSZ + MTX and MTX alone&lt;sup&gt;4, 21-25&lt;/sup&gt;</td>
<td>Low for trials</td>
<td>Downgraded because open label design; high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>SSZ + MTX and MTX alone&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Insufficient for obs evidence</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>SSZ compared with MTX, with or without concomitant PNL&lt;sup&gt;27, 28&lt;/sup&gt;</td>
<td>Insufficient for both trials and obs evidence</td>
<td>Trials: Downgraded because high attrition; large baseline differences between groups; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size Obs evidence: Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>TOF compared with MTX alone&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>SSZ compared with MTX, with concomitant PNL&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; direction of effect varies; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>TOF compared with MTX alone&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large CIs; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>SSZ compared with MTX, with or without concomitant PNL&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large baseline differences between groups; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>SSZ + MTX and MTX alone&lt;sup&gt;4, 21, 22, 24, 25&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>TOF compared with MTX&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>No significant difference between SSZ + MTX and MTX alone&lt;sup&gt;4, 21-25&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>No significant difference between PNL + MTX + SSZ + HCQ and monotherapy with MTX or SSZ&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>SSZ compared with MTX, with or without concomitant PNL&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Insufficient for both trials and obs evidence</td>
<td>Trials: Downgraded because high attrition; not enough events to meet optimal information size; and large baseline differences between groups</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>TOF compared with MTX alone&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences between SSZ + MTX and MTX alone&lt;sup&gt;4, 21-25&lt;/sup&gt;</td>
<td>Low for trials</td>
<td>Downgraded because open label design; high attrition; and imprecision</td>
</tr>
<tr>
<td>Harms</td>
<td>D/C attributable to AEs</td>
<td>SSZ + MTX and MTX alone&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Insufficient for obs evidence</td>
<td>Downgraded because of high risk of selection bias for treatment discontinuation; high risk of confounding by indication; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>TOF compared with MTX&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Harms</td>
<td>D/C attributable to AEs</td>
<td>SSZ compared with MTX, with or without concomitant PNL&lt;sup&gt;23, 28&lt;/sup&gt;</td>
<td>Insufficient for both trials and obs evidence</td>
<td>Trials: Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
</tbody>
</table>

<sup>a</sup> Response defined by ACR or DAS28

ACR = American College of Rheumatology; AE = adverse event; CI = confidence interval; csDMARD = conventional synthetic disease modifying anti-rheumatic drug; D/C = discontinuation; DAS28 = Disease Activity Score based on 28 joints; HCQ = hydroxychloroquine; MTX = methotrexate; N/A = not applicable; Obs = observational; PNL = prednisolone; PRED = prednisone; SAE = serious adverse event; SSZ = sulfasalazine; TOF = tofacitinib; tsDMARD = targeted synthetic DMARD.
Table C. Benefits and harms of biologic DMARDs for early RA treatment

<table>
<thead>
<tr>
<th>Outcome Type</th>
<th>Specific Outcome</th>
<th>Results</th>
<th>Strength of Evidence</th>
<th>Summary of Rationale for Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease activity</td>
<td>Response</td>
<td>No significant difference between non-TNF biologic (ABA) + MTX and ABA alone&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response for TNF biologic (ADA) + MTX than ADA alone and for ADA than MTX&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response for TNF biologic (ETN) + MTX than MTX alone&lt;sup&gt;12, 14&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response for non-TNF biologics (ABA or RIT) + MTX than MTX alone&lt;sup&gt;7, 30, 31&lt;/sup&gt;</td>
<td>Moderate</td>
<td>ABA + MTX: Downgraded because high attrition; RIT + MTX: Not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>Significantly higher remission for TNF biologic (ADA) + MTX than ADA alone and for ADA than MTX&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for TNF biologic (ADA) + MTX than ADA alone and for ADA than MTX&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for non-TNF biologic (TCZ) + MTX than TCZ alone and for TCZ than MTX&lt;sup&gt;32, 33&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for TNF biologic (ETN) alone and combined with MTX than MTX alone&lt;sup&gt;12, 14&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for non-TNF biologic (TCZ) + MTX than MTX alone&lt;sup&gt;12, 14&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for non-TNF biologic (RIT) + MTX than MTX alone&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response with TNF biologic (ADA) + MTX compared with MTX alone&lt;sup&gt;15, 16, 34-37&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response with TNF biologic (CZP) + MTX compared with MTX alone&lt;sup&gt;38, 39&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition; large confidence intervals; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Significantly improved response with TNF biologic (IFX) + MTX than csDMARD combination therapy&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response&lt;sup&gt;a&lt;/sup&gt;</td>
<td>No significant difference between TNF biologic (IFX) + MTX and csDMARD combination and csDMARD combination therapies&lt;sup&gt;60&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>No significant difference between non-TNF biologic (ABA) + MTX and ABA alone&lt;sup&gt;7&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>Significantly higher remission for non-TNF biologic (TCZ) + MTX than TCZ alone and for TCZ than MTX(^{32,33})</td>
<td>Low</td>
<td>Downgraded because large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>Significantly increased remission for TNF biologics (ADA, CZP, ETN, IFX) plus MTX, or TNF biologic alone (ETN), compared with MTX alone(^{12-17,34-37,41})</td>
<td>Low(^b)</td>
<td>ADA + MTX: Downgraded because high attrition; CZP + MTX: Downgraded because high attrition; large CIs; and not enough events to meet optimal information size; ETN + MTX and ETN alone: Downgraded because medium level of study limitations, and not enough events to meet optimal information size; IFX + MTX: Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>No significant difference between TNF biologic (IFX) + csDMARD combination and csDMARD combination therapies(^{10})</td>
<td>Low</td>
<td>Downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>Significantly higher remission for non-TNF biologic (TCZ) + MTX than MTX alone(^{32,33})</td>
<td>Low</td>
<td>Downgraded because medium level of study limitations, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>Significantly less radiographic progression for some biologics (TNF: ADA, CZP; non-TNF: ABA) plus MTX compared with MTX alone(^{13,15,31})</td>
<td>Low(^b)</td>
<td>ADA + MTX: Downgraded because high attrition, and large CIs cross appreciable benefits or harms; CZP + MTX: Downgraded because high attrition; large CIs; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>No significant difference between TNF biologic (IFX) + csDMARD combination therapy compared with csDMARD combination therapy alone(^{40})</td>
<td>Low</td>
<td>Downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response(^a)</td>
<td>Non-TNF biologic (TCZ) + MTX compared with TCZ alone and TCZ compared with MTX(^{12,31})</td>
<td>Insufficient</td>
<td>Downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response⁹</td>
<td>IFX + MTX compared with MTX alone¹⁷, ¹⁸, ⁴¹</td>
<td>Insufficient</td>
<td>Downgraded because not enough events to meet optimal information size; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response⁹</td>
<td>TNF biologic (ADA or ETN) compared with non-TNF biologic (RIT)⁸</td>
<td>Insufficient</td>
<td>Downgraded because no ITT analysis; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response⁹</td>
<td>TCZ + MTX compared with MTX alone³², ³³</td>
<td>Insufficient</td>
<td>Downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Response⁹</td>
<td>ADA + MTX compared with csDMARD combination with PRED⁹</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; not enough events to meet optimal information size; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>ADA + MTX compared with csDMARD combination with PRED⁹</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; not enough events to meet optimal information size; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Remission</td>
<td>TNF biologic (ADA or ETN) compared with non-TNF biologic (RIT)⁸</td>
<td>Insufficient</td>
<td>Downgraded because no ITT analysis; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>IFX + MTX compared with MTX alone¹⁷, ⁴¹</td>
<td>Insufficient</td>
<td>Downgraded because not enough events to meet optimal information size; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Disease activity</td>
<td>Radiographic Changes</td>
<td>ADA + MTX compared with csDMARD combination with PRED⁹</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; not enough events to meet optimal information size; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Significantly greater improvement in TNF biologic (ADA) plus MTX than ADA alone and for ADA than MTX¹⁵</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Significantly greater improvement for TNF biologic (ADA) + MTX than MTX alone¹⁵, ¹⁶, ³⁴, ³⁷</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Significantly greater improvement for non-TNF biologic (RIT) combined with MTX than MTX alone³⁰</td>
<td>Moderate</td>
<td>Downgraded because not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>No significant differences in functional capacity for ABA + MTX vs. ABA Low or for ABA vs. MTX⁷</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Significantly greater improvement in TNF biologic (CZP, IFX) plus MTX than MTX alone&lt;sup&gt;13, 17, 41&lt;/sup&gt;</td>
<td>Low&lt;sup&gt;6&lt;/sup&gt;</td>
<td>CZP + MTX: Downgraded because high attrition; large confidence intervals; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Significantly greater improvement in non-TNF biologic (RIT) than TNF biologics (ADA, ETN)&lt;sup&gt;8&lt;/sup&gt;</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>Mixed results for TNF biologic (ETN) or non-TNF biologic (ABA) plus MTX compared with MTX alone&lt;sup&gt;7, 12, 14, 31&lt;/sup&gt;</td>
<td>Low&lt;sup&gt;6&lt;/sup&gt;</td>
<td>ABA + MTX: Downgraded because high attrition; ETN + MTX: Downgraded because direction of effect varies, and large CIs</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>No significant difference between TNF biologic (IFX) + csDMARD combination and csDMARD combination therapies&lt;sup&gt;60&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>TCZ + MTX vs. TCZ and TCZ vs. MTX&lt;sup&gt;12, 33&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>ADA + MTX compared with csDMARD combination with PRED&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Functional Capacity</td>
<td>N/A</td>
<td>TCZ + MTX compared with MTX&lt;sup&gt;12, 33&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences between TNF biologic (ADA) + MTX and ADA alone or between ADA and MTX&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences between non-TNF biologic (TCZ) + MTX and TCZ alone or between TCZ and MTX&lt;sup&gt;12, 33&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences between non-TNF biologic (TCZ) + MTX and MTX alone&lt;sup&gt;12, 33&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences among non-TNF biologic (RIT) + MTX and MTX alone&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Moderate</td>
<td>Downgraded because single-study body of evidence</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>No significant differences between non-TNF biologic (ABA) + MTX and ABA alone or between ABA and MTX&lt;sup&gt;7&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributed to AEs</td>
<td>No significant differences between TNF biologics (ADA, CZP, ETN, IFX) plus MTX and MTX alone&lt;sup&gt;12-17, 34-37&lt;/sup&gt;</td>
<td>Low&lt;sup&gt;b&lt;/sup&gt;</td>
<td>ADA + MTX: Downgraded because high attrition; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributed to AEs</td>
<td>No significant difference between non-TNF biologic (ABA) plus MTX and MTX alone&lt;sup&gt;31&lt;/sup&gt;</td>
<td>Low</td>
<td>Downgraded because high attrition</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributed to AEs</td>
<td>No significant difference between TNF biologic (IFX) + MTX or IFX + csDMARD combination and csDMARD combination therapies&lt;sup&gt;10, 40&lt;/sup&gt;</td>
<td>Low&lt;sup&gt;b&lt;/sup&gt;</td>
<td>IFX + MTX: Downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs</td>
<td>ADA + MTX compared with csDMARD combination with PRED&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Insufficient</td>
<td>Downgraded because high attrition; not enough events to meet optimal information size; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Outcome Type</td>
<td>Specific Outcome</td>
<td>Results</td>
<td>Strength of Evidence</td>
<td>Summary of Rationale for Strength of Evidence</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Harms</td>
<td>SAEs and D/C attributable to AEs</td>
<td>TNF biologic (ADA or ETN) compared with non-TNF biologic (RIT)(^a)</td>
<td>Insufficient</td>
<td>Downgraded because no ITT analysis; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
</tbody>
</table>

\(^a\) Response defined by ACR or DAS28.\(^b\) Strength of evidence grade applies to each specific drug therapy named in the Results column.

ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; AE = adverse event; CI = confidence interval; csDMARDs = conventional synthetic disease-modifying antirheumatic drug; CZP = certolizumab pegol; D/C = discontinuation; DAS28 = Disease Activity Score based on 28 joints; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; IFX = infliximab; ITT = intent-to-treat; MTX = methotrexate; N/A = not applicable; PRED = prednisone; RIT = rituximab; SAE = serious adverse event; TCZ = tocilizumab; TNF = tumor necrosis factor.
The combinations of several TNFs (ADA, CZP, IFX) and non-TNF biologics (rituximab) plus MTX also produced greater functional capacity than MTX monotherapy (all moderate or low SOE). The results for the remainder of the biologics (ETN, ABA, TCZ) were inconclusive. IFX (TNF) plus MTX, when compared with csDMARD combination therapy, resulted in reduced disease activity, but the groups did not differ with respect to other outcomes. Likewise, when IFX was combined with multiple csDMARDs (MTX + SSZ + hydroxychloroquine) and prednisone (PRED) and compared with csDMARD combination therapies plus PRED, outcomes did not differ. No data are available for IFX monotherapy; it is approved and generally given in combination with MTX.

NWMA found higher overall discontinuation rates for MTX monotherapy than combination therapy consisting of biologic DMARDs (ADA, CZP, ETN) plus MTX (range of RR, 1.52 [95% CI, 1.02 to 2.27] to 1.77 [95% CI, 1.32 to 2.36]). However, neither serious adverse events nor discontinuations attributable to adverse events differed between the groups (low SOE). Lack of efficacy is a possible reason that patients may have discontinued the therapy or withdrawn from these studies. Evidence was insufficient for drawing conclusions about several other drug therapy combinations or head-to-head comparisons.

Combination Therapies and Treatment Strategies: Four RCTs evaluated different combination therapies and treatment strategies for early RA with moderate to high disease activity; in addition, two observational studies contributed data on harms. Patients receiving combination therapy containing MTX, SSZ plus tapered high-dose PRED (60 mg/day tapered to 7.5 mg/day), or MTX plus IFX (TNF biologic) had lower disease activity (moderate SOE) and greater functional capacity (low SOE) at 1 year (DAS<2.4: 71 to 74% vs. 53 to 64%) and less radiographic progression (moderate SOE) at 4 and 5 years (modified Total Sharp/van der Heijde score [mTSS]: 2.5 to 3.0 vs. 5.0 to 5.5) than patients receiving sequential csDMARD or step-up combination therapies starting with MTX. Groups did not differ with respect to remission (moderate SOE), serious adverse events (low SOE), or other outcomes over the longer term. We could not draw any conclusions about immediate or step-up combination therapies containing MTX and either additional csDMARD(s) or ETN (TNF biologic).

Results Among Subgroups of Patients: Only four RCTs compared drug therapies among different subpopulations defined by demographics, disease activity, or coexisting conditions (Table A). We could not draw any conclusions about response rates between older and younger patients or about response rate and radiographic changes between people with different levels of disease activity who were taking MTX with or without a TNF biologic (ADA or IFX). Evidence was also insufficient to draw any conclusions about serious adverse events as defined by FDA between older and younger patients who were taking MTX or the TNF biologic ETN. No data were available for the other agents.

Discussion and Findings in Context

We conducted a systematic review and NWMA to update the 2012 review of the comparative effectiveness of drug therapies for rheumatoid arthritis (RA);1 in this report we focused solely on early RA in adults (within 1 year of diagnosis). Although level of disease severity was not used as a criterion to determine study eligibility, all of our early RA studies included patients with moderate to high disease activity. In a clinical setting, patients with early RA may present with varying levels of severity.

Current clinical practice guideline recommendations for therapy for patients with early RA and moderate-to-high disease severity are consistent with our findings but ours go further and
also support additional therapies in patients with moderate to high disease activity.\textsuperscript{42, 43} When disease activity remains moderate or high (DAS28 ≥ 3.2) despite initial treatment, the ACR RA guidelines recommend double or triple csDMARD therapy or a TNF or non-TNF biologic (with or without MTX). We found that when biologics were used in combination with MTX therapy, patients achieved better disease control, higher functional capacity, and higher remission rates than with biologics or MTX monotherapy.

This report assessed the comparative effectiveness based on current evidence. While not directly comparable, the ACR clinical guidelines move beyond evidence to make recommendations when evidence is limited. Clinical practice guidelines use systematic reviews as evidence and if evidence is not enough they may consider other resources. The recommendations were based on a consideration of the balance of relative benefits and harms of the treatment options under consideration and the quality of the evidence. Additionally, the ACR recommendations included consideration of patients’ values and preferences.

Although the literature in this review supports the effectiveness of MTX plus biologics in early RA for patients with medium and high disease activity, it is not currently the standard of care for a number of potential reasons. Some data indicate that certain patients with early disease may respond well to MTX monotherapy, although no information is available about how to predict which patients will do so. Second, many insurers require inadequate response to MTX as a prerequisite to adding a biologic (this policy is probably based on findings of the effectiveness of MTX). Third, patients may be wary, for a variety of reasons, of a combination therapy approach in early disease (e.g., cost, side effects, injections). Additionally, patients may find it difficult to balance the burden of multiple drugs and potentially higher risks.

Current European League Against Rheumatism early RA guidelines recommend adding a TNF or non-TNF biologic to a csDMARD but only when patients have poor prognostic factors (e.g., high disease activity, early joint damage, autoantibody positivity).\textsuperscript{43} The available evidence in this review (from 10 studies comparing combinations of biologics and MTX with either biologic or MTX monotherapy) supports this recommendation. Specifically, these studies indicate that patients receiving combination therapies may achieve higher remission rates than those receiving monotherapy.\textsuperscript{12-15, 17, 32-34, 37, 41}

However, our data were limited because we did not find available studies that specifically examined therapies in patients with early RA and less severe disease activity compared with patients with early RA plus poor prognostic factors.

**Contextual Questions:** In one review comparing early versus delayed treatment trials, RA patients treated immediately at presentation with csDMARDs had improved patient function and reduced radiographic progression than patients whose DMARD treatment had been delayed 6 to 12 months.\textsuperscript{44} Some of the barriers preventing early RA patients from accessing indicated drug therapies included access to primary health care services, difficulties in diagnosing RA in the primary care setting, obtaining of insurer approval of biologic DMARDs, high out-of-pocket expenses, and limited access to specialty care, especially in rural areas.\textsuperscript{45} Other challenges included contraindications for some drug therapies, especially among patients with coexisting conditions and older patients, and patients’ reluctance to begin therapies.\textsuperscript{46, 47}
Key Limitations and Research Gaps

Limitations of the Evidence Base

We encountered a limited number, or a complete lack, of trials or studies about some drugs (or entire drug classes) on early RA patients. These gaps in the evidence base prohibited us from conducting an even more comprehensive evaluation and synthesis. Specifically, we found no eligible trials or other studies for biosimilar drugs and sarilumab, although FDA approved them for use among early RA patients within the past 5 years. We also found only limited evidence for tsDMARDs. In both cases, we assume that more evidence will emerge in coming years.

Information about harms from the included studies was scarce. This report includes information related to rates of serious adverse events or numbers or rates of patient discontinuations attributed to adverse events. However, we found little or no information about more common side effects that are likely important to patients. This information is widely available in the prescribing information and is based on data from the registration trials. However, most of the time it is not included in the publication.

In addition, the important corticosteroid and MTX comparisons were from studies that used different, or variable, dosage ranges. This made quantitative synthesis (i.e., meta-analyses) difficult if not impossible for these drugs.

Moreover, the population of interest was confined to patients with early RA (1 year or less). Some debate remains as to whether “early RA” should include patients diagnosed with RA within the previous 2 years (rather than 1 year). Given this variability, a European task force of experts in RA and clinical trial methodology recommended defining early RA as no more than 1 year of diagnosed disease duration. Defining early RA this way subsumes the ACR definition of duration as less than 6 months of disease symptoms, but it is consistent with early RA in clinical rheumatology practice. Our search excluded 7 studies (reported in 10 articles) with RA from 1 to 2 years’ duration. On brief review of the 7 studies, findings did not differ from the current report. Additional evidence on treatment comparisons might be gained by expanding the definition to 2 years, but the more clinically rigorous 1-year specification is in line with current practice.

Finally, because of the lack of data for some therapies, this update will itself need to be updated when more and better trials are published. Specifically, a future update may include data from newer drugs currently under review, the biosimilars, longer trials, and more information on harms.

Research Gaps

Future studies need to compare therapy strategies in patients diagnosed with early RA who have different degrees of disease activity or poor prognostic factors and what, if any, therapies patients have already tried. Documenting these types of variables at baseline may provide important insights into the impact of the full range of treatment options on this early RA subgroup. Additionally, the evidence base will improve as studies begin to use a consistent definition of early RA.

Information is needed about the performance of drugs in subpopulations of patients defined by various important characteristics. These characteristics include health status, sociodemographic variables such as age or race and ethnicity, and coexisting conditions, particularly chronic conditions that occur commonly in patients with RA (such as diabetes).
Finding ways to study these patient subgroups is crucial if research is to help clinicians select appropriate treatments for such populations.

Finally, for early RA patients, we need longer term data to assess the overall impact of medications that we know may be beneficial initially, but we do not know their effectiveness over time. Thus, trials with long treatment periods (5 or more years) and even posttreatment followup are needed. These longer trials can provide more and better information on important outcomes such as remission, recurrence, and quality of life; adherence to potentially complex medical regimens; and mild, moderate, and severe adverse events. Longer trials would also yield insights into whether starting with a biologic in early disease improves the long-term prognosis of RA.
Introduction

Condition

Rheumatoid arthritis (RA) is an autoimmune systemic inflammatory arthritis. RA affects 1 percent of the world’s population, including more than 1 million American adults. RA is characterized by synovial inflammation of joints, which can lead to progressive erosion of bone, irreversible damage to the joint, loss of function, and resultant disability. The average incidence of RA in the United States is approximately 70 per 100,000 adults annually. RA can develop at any age, but incidence increases with age, peaking in the fifth decade. The incidence of RA is 2 to 3 times higher in women.

Etiology

The etiology of RA is incompletely understood, but multiple environmental and genetic factors contribute to the development of the disease. Obesity, smoking, and nulliparity increase the risk. Other environmental risk factors associated with RA, although not well understood, include low socioeconomic status and viral and bacterial infections, including those caused by periodontal and lung pathogens. Additionally, researchers using animal models are investigating the contribution of the microbiome to the development of RA. Rates of RA development are higher in monozygotic twins, implicating genetics as a contributing factor. Genome-wide association studies have characterized more than 100 loci associated with RA risk; most involve immune mechanisms. The driving genetic force is the MHC (Major Histocompatibility Complex) Class II shared epitope. The confluence of both environmental and genetic factors in individuals (epigenetics) also contributes to the pathogenesis of RA.

Burden of Disease

Disability associated with RA is significant. More than 35 percent of patients with RA have a work disability after 10 years. The lifespan of RA patients is 3 to 12 years shorter than that of the general population. Patients with RA, especially those with high disease activity, are at increased risk of cardiovascular disease, which contributes to higher mortality risk.

Definitions of Early RA and Challenges With the Definitions

Defining RA for the purposes of this systematic review proved to be most challenging, because no consensus exists on the definition of early RA. As our knowledge of the pathogenesis of RA has advanced, there is a broad understanding that RA begins well before the development of the well-characterized clinical signs and symptoms of joint stiffness, pain, and swelling. The disease exists on a continuum from this preclinical stage to established disease involving the typical inflammatory disease with damage to the joints.

Definitions, by expert groups, of the beginning of early RA include symptom onset to when a clinician diagnoses RA. Experts base their initial treatment recommendations on either time from diagnosis or, more stringently, time from initial symptoms. In terms of duration, the American College of Rheumatology (ACR) defines early RA as the first 6 months of symptoms, while other organizations advocate for up to 2 years after diagnosis.

Note: The reference list follows the appendixes.
In theory, treating RA early, prior to joint damage, leads to better outcomes overall than treating disease later in the course. In addition, there is increasing interest in evaluating the effectiveness of therapy in patients thought to be at high risk of developing RA but who do not yet meet the ACR/European League Against Rheumatism (EULAR) classification criteria. This is a compelling idea that presses the question of whether we can prevent the development of full-blown RA in this subset of patients. In addition, it resets the notion of how we will define “early” disease going forward.

The course of RA is highly variable; this factor precludes using a specific biological or physical benchmark or marker to identify those with early RA. For example, some researchers have suggested that early RA should be defined as the time period before patients develop bone erosion, but some patients never develop erosions.

Given this variability, a European task force of experts in RA and clinical trial methodology recommended defining early RA as no more than 1 year of diagnosed disease duration. Defining early RA this way subsumes the ACR definition of duration as less than 6 months of disease symptoms, but it is consistent with early RA in clinical rheumatology practice. Given the above caveat and limitations of placing of boundaries on the continuum of early RA, this is the basic definition (no more than 1 year of diagnosed RA) we adopted for this systematic review update.

The goal of separating early disease from late disease, however one defines these stages, is not to assess whether, or imply that, response to certain therapeutics differs by stage of disease but to provide a framework to facilitate the discussion about the effects of treating RA earlier rather than later.

**Current Practice and Treatment Strategies**

In all patients with early RA, experts recommend early treatment with the goal of sustained remission or low disease activity. RA treatment aims to control pain and inflammation and, ultimately, slow the progression of joint destruction and disability. Disease activity, categorized as low, moderate, and high by validated scales, can guide the initial choice and subsequent adjustment of therapy including any disease-modifying antirheumatic drug (DMARD) adjustment. Disease activity, functional assessment, patient-reported outcomes, and structural damage observed on radiographs should be measured regularly. Based on these measurements, clinicians should assess drug therapy at regular intervals until patients reach the treatment target, which is ideally remission.

For symptomatic early RA, the ACR recommends a treat-to-target approach to achieving remission or low disease activity, rather than a nontargeted approach; this guidance is based on low strength of evidence. Treating to target includes regularly monitoring disease activity and adverse events and escalating treatment if patients do not reach a treatment target (ideally remission). DMARD monotherapy, methotrexate (MTX) preferred, is initially recommended instead of double or triple therapy in patients who have never taken a DMARD (low strength of evidence). If disease activity remains moderate or high, using double or triple combination DMARDs or adding a tumor necrosis factor (TNF) or non-TNF biologic DMARD is recommended (low strength of evidence). Low-dose glucocorticoids (≤10 mg/day prednisone or equivalent) are recommended in addition if disease activity is moderate or high despite DMARD use (low to moderate strength of evidence).

The EULAR task force recommends starting treatment with DMARDs as soon as the RA diagnosis is made. It also recommends a treat-to-target approach to achieve remission or low disease activity. EULAR advocates using conventional synthetic DMARDs (csDMARDs) as
monotherapy or combination therapy for the initial DMARD treatment strategy. The csDMARDs include hydroxychloroquine (HCQ), leflunomide (LEF), MTX, and sulfasalazine (SSZ). If patients who do not have poor prognostic factors such as high disease activity, early joint damage, or autoantibody positivity do not achieve the treatment target with the first DMARD, such as MTX, then clinicians should consider using a different csDMARD (e.g., LEF or SSZ). If patients do have poor prognostic factors, then adding a TNF or non-TNF biologic to the first DMARD is recommended.

The EULAR task force regards all currently approved biologic DMARDs as similarly effective and similarly safe after csDMARD failure. Anakinra is the exception, as it has not shown strong efficacy when compared with other DMARDs. The ACR guidelines also did not include anakinra because of its infrequent use in RA and the lack of new data on it since 2012.

**Drugs Approved by the U.S. Food and Drug Administration**

Available therapies for RA include corticosteroids, csDMARDs, TNF and non-TNF biologics, targeted synthetic DMARDs (tsDMARDs), and biosimilars. Table 1 provides the names of specific pharmaceutical agents in these categories; it is ordered roughly from oldest to newest drugs in terms of approvals by the U.S. Food and Drug Administration (FDA).

<table>
<thead>
<tr>
<th>Group</th>
<th>Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroids</td>
<td>Methylprednisolone, prednisone, prednisolone</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>Hydroxychloroquine, leflunomide, methotrexate, sulfasalazine</td>
</tr>
<tr>
<td>TNF biologics</td>
<td>Adalimumab, certolizumab pegol, etanercept, golimumab, infliximab</td>
</tr>
<tr>
<td>Non-TNF biologics</td>
<td>Abatacept, rituximab, tocilizumab, sarilumab^a</td>
</tr>
<tr>
<td>tsDMARDs</td>
<td>Tofacitinib^a</td>
</tr>
<tr>
<td>Biosimilars</td>
<td>Adalimumab-atto,^a infliximab-dyyb,^a infliximab-abda,^a etanercept-szzs^a</td>
</tr>
</tbody>
</table>

^a New medications that the U.S. Food and Drug Administration has approved since 2012.

csDMARD = conventional synthetic disease-modifying antirheumatic drug; TNF = tumor necrosis factor; tsDMARD = targeted synthetic disease-modifying antirheumatic drug.

**Challenges in Treating Early Rheumatoid Arthritis**

Challenges and controversies related to early RA include several main issues. The first issue surrounds the role of newly approved drugs in the treatment strategies in the context of older medications. The number of drugs for treating early RA continues to increase with the addition of tsDMARDs, newer biologics, and biosimilars. It is important to examine whether additional improvement in patient outcomes is gained and if improvements are tempered by potential harms. A second issue is the appropriate use and order or combination of different therapeutic options. There is a dizzying array of RA medications, and combining them and designing treatment strategies demand additional choices from clinicians. Finally, identifying the optimal approach to managing RA therapy in the context of coexisting conditions (e.g., malignancy, infections like hepatitis C, congestive heart failure, diabetes) is a third challenge; pregnancy can also be an issue. Careful consideration of RA treatment drug choice is essential in these populations.

Clinicians face the challenge of identifying which DMARD to initiate for patients with early RA. Traditionally, biologics are not approved as first-line treatment. Nevertheless, clinicians still
must decide whether to institute csDMARDs, tsDMARDs, or biologics earlier in the disease course.

The overarching principle should be to treat to target using disease activity remission criteria. Among the questions clinicians have are whether they should adopt one of the following approaches:

1. Apply step-up treatment (i.e., progress from single therapy to combination therapy) or
2. Apply step-down therapy (i.e., begin with combination therapy and back down treatment when symptoms are under control).

Treatment tapering or stopping strategies are also debated. When patients respond (e.g., reach low disease activity) or reach remission, the main question is whether DMARDs can be tapered or stopped. This quandary raises questions about other issues, such as how to define remission or set the appropriate taper. Also, patients may want to taper off DMARDs when their symptoms have improved; however, clinically, inflammation may be ongoing, rendering tapering off potentially inappropriate.

Scope and Key Questions

Scope of the Review

This systematic review and meta-analysis updates the 2012 report, Drug Therapy for Rheumatoid Arthritis in Adults: An Update. However, the targeted scope for this review focuses solely on patients with early RA.

Evidence Gaps From Prior Review

In the 2012 review, the existing evidence was insufficient to draw conclusions on the best treatment regimen for patients with early RA. Mainly, studies were of limited duration. This factor did not allow comparisons of whether early initiation of a biologic in addition to a csDMARD improved disease activity, radiographic findings, functional capacity, or quality of life compared with csDMARDs (HCQ, LEF, MTX, SSZ). No studies investigated efficacy, effectiveness, and harms among subgroup populations.

New Therapies

Since the 2012 review, information from additional clinical trials of four biosimilar drugs, a tsDMARD (an oral synthetic Janus kinase inhibitor), and one non-TNF biologic have become available. In addition, studies continue to be published on established therapies.

Newly approved drugs are marked with a footnote in Table 1 above. Few data are available on the efficacy of these drugs; even less is known about the effectiveness and harms compared with those of the previously existing drugs. Only a few large head-to-head trials have been conducted on any of the existing medications or new therapies. Consequently, examining the current literature as to whether all drugs examined have longer follow-up periods or include subgroups would be important knowledge gained in this review.

What This Review Aims To Do

This review focuses on patients with early RA as defined earlier. It updates the 2012 review on the comparative effectiveness of drug therapies with respect to disease activity, joint damage,
patient-reported symptoms, functional capacity, and quality of life. We also examine comparative harms of drug therapies in terms of tolerability, adherence, and adverse effects. Finally, we examine comparative effectiveness and harms of drug therapies in patient subgroups. We address four Key Questions (KQ) and two Contextual Questions (CQs).

Key Questions

**KQ 1:** For patients with early RA, do drug therapies differ in their ability to reduce disease activity, slow or limit the progression of radiographic joint damage, or induce remission?

**KQ 2:** For patients with early RA, do drug therapies differ in their ability to improve patient-reported symptoms, functional capacity, or quality of life?

**KQ 3:** For patients with early RA, do drug therapies differ in harms, tolerability, patient adherence, or adverse effects?

**KQ 4:** What are the comparative benefits and harms of drug therapies for early RA in subgroups of patients based on disease activity, prior therapy, demographics (e.g., women in their childbearing years), concomitant therapies, and presence of other serious conditions?

Contextual Questions

CQs are not systematically reviewed. Rather, we use evidence readily available to us from our literature searches for the KQs and additional searches as needed.

**CQ 1:** Does treatment of early RA improve disease trajectory and disease outcomes compared with the trajectory or outcomes of treatment of established RA?

**CQ 2:** What barriers prevent individuals with early RA from obtaining access to indicated drug therapies?

Analytic Framework

Figure 1 visually depicts the KQs within the context of the PICOTS (Populations, Interventions, Comparators, Outcomes, Timing, Setting, Study design) or eligibility criteria described in detail in the Methods section below. In general, the figure illustrates the potential outcomes that adults with early rheumatoid arthritis may experience following treatment with corticosteroids, csDMARDs, TNF biologics, non-TNF biologics, or biosimilars versus any of these same treatments. KQ 1 considers whether patients may experience benefits in intermediate outcomes such as disease activity, joint damage, and remission. KQ 2 asks the same question, but regarding benefits in final health outcomes such as functional capacity, quality of life, and patient-reported symptoms. KQ 3 considers the potential adverse effects of the medications.
described previously. KQ 4 addresses the comparative benefits and harms of medications for subgroups of patients with early rheumatoid arthritis.

**Figure 1. Analytic framework for drug therapy for early rheumatoid arthritis**

![Analytic framework for drug therapy for early rheumatoid arthritis](image)

KQs 1, 4

**Intermediate outcomes**
- Disease activity
- Joint damage
- Remission

**Final health outcomes**
- Functional capacity
- Quality of life
- Patient-reported symptoms

KQs 2, 4

Corticosteroids; csDMARDs; TNF biologics; non-TNF biologics; tsDMARDs; biosimilars

Adults with early RA

Adverse effects of treatment

csDMARD = conventional synthetic disease-modifying antirheumatic drug; KQ = Key Question; RA = rheumatoid arthritis; TNF = tumor necrosis factor; tsDMARD = targeted synthetic disease-modifying antirheumatic drug.

**Organization of This Report**

We describe our methods next and then present our key findings in the Results chapter. In the Discussion chapter, we explore the implications of our findings and examine the limitations of the evidence base in this review, clarify gaps in the knowledge base, and offer recommendations for future research. References follow the appendixes. The main report has several appendixes, as follows: Appendix A, Search Strings; Appendix B, Excluded Articles; Appendix C, Detailed Evidence Table; Appendix D, Risk of Bias Ratings and Rationales for Included Studies; Appendix E, Strength of Evidence for Key Question 1–4 Outcomes; Appendix F, Eligible Clinical and Self-Reported Scales and Instruments Commonly Used in Eligible Studies of Drug Therapy for Rheumatoid Arthritis; Appendix G, Tests of Consistency for Main Network Meta-Analyses; Appendix H, Supplementary Primary Network Meta-Analyses; Appendix I, Sensitivity Analyses for Network Meta-Analyses; Appendix J, Expert Guidance and Review; and Appendix K, PCORI Methodology Standards Checklist.
Methods

The methods for this systematic review (SR) follow the Agency for Healthcare Quality and Research (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (available at [http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm](http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm)) and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist. The main sections in this chapter reflect the elements of the protocol established for this review of treatments of patients with early rheumatoid arthritis (RA). The final protocol can be found on the Effective Health Care Web site ([https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/research-protocol/](https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/research-protocol/)); it is also registered on PROSPERO (available at [http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017079260](http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017079260)). All methods and analyses were determined a priori.

Stakeholders, including Key Informants and Technical Experts, participated in a virtual workshop facilitated by Patient-Centered Outcomes Research Institute (PCORI) in December 2016 to help formulate the research protocol (further details in Appendix J). Key Informants in the workshop included end users of research, such as patients and caregivers; practicing clinicians; relevant professional and consumer organizations; purchasers of health care; and others with experience in making health care decisions. Technical Experts in the workshop included multidisciplinary groups of clinical, content, and methodological experts who provided input in defining populations, interventions, comparisons, and outcomes and identified particular studies or databases to search. They were selected to provide broad expertise and perspectives specific to drug therapy for RA in adults.

Criteria for Inclusion/Exclusion of Studies in the Review

The criteria for inclusion and exclusion of studies are designed to identify research that can answer the four Key Questions (KQs) concerning early RA specified in the introduction. The criteria are based on the population, intervention/exposure, comparator, outcomes, time frames, country and clinical settings, and study design (PICOTS) shown in Table 2.

<table>
<thead>
<tr>
<th>PICOTS</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>All KQs: Adult outpatients 19 years of age or older with a diagnosis of early RA, defined as 1 year or less from disease diagnosis; we will include studies with mixed populations if &gt;50% of study populations had an early RA diagnosis. KQ 4 only: Subpopulations by age, sex or gender, race or ethnicity, disease activity, prior therapies, concomitant therapies, and other serious conditions</td>
<td>Adolescents and adult patients with disease greater than 1 year from diagnosis; inpatients</td>
</tr>
<tr>
<td>Intervention</td>
<td><strong>Corticosteroids:</strong> methylprednisolone, prednisone, prednisolone <strong>csDMARDs:</strong> hydroxychloroquine, leflunomide, methotrexate, sulfasalazine <strong>TNF biologics:</strong> adalimumab, certolizumab pegol, etanercept, golimumab, infliximab <strong>Non-TNF biologics:</strong> abatacept, rituximab, sarilumab, tocilizumab <strong>tsDMARDs:</strong> tofacitinib <strong>Biosimilars:</strong> adalimumab-atto, infliximab-dyyb, infliximab-abda, etanercept-szsz</td>
<td>Anakinra is excluded because, although it is approved for RA, clinically it is not used for this population</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PICOTS</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>All KQs: Adult outpatients 19 years of age or older with a diagnosis of early RA, defined as 1 year or less from disease diagnosis; we will include studies with mixed populations if &gt;50% of study populations had an early RA diagnosis. KQ 4 only: Subpopulations by age, sex or gender, race or ethnicity, disease activity, prior therapies, concomitant therapies, and other serious conditions</td>
<td>Adolescents and adult patients with disease greater than 1 year from diagnosis; inpatients</td>
</tr>
<tr>
<td>Intervention</td>
<td><strong>Corticosteroids:</strong> methylprednisolone, prednisone, prednisolone <strong>csDMARDs:</strong> hydroxychloroquine, leflunomide, methotrexate, sulfasalazine <strong>TNF biologics:</strong> adalimumab, certolizumab pegol, etanercept, golimumab, infliximab <strong>Non-TNF biologics:</strong> abatacept, rituximab, sarilumab, tocilizumab <strong>tsDMARDs:</strong> tofacitinib <strong>Biosimilars:</strong> adalimumab-atto, infliximab-dyyb, infliximab-abda, etanercept-szsz</td>
<td>Anakinra is excluded because, although it is approved for RA, clinically it is not used for this population</td>
</tr>
</tbody>
</table>

Table 2. Eligibility criteria for review of treatments for early rheumatoid arthritis
PICOTS | Inclusion | Exclusion
--- | --- | ---
Comparator<br>For head-to-head RCTs, head-to-head nRCTs, and prospective, controlled cohort studies (all KQs): Any active intervention listed above<br>For additional observational studies of harms (i.e., overall KQ 3 and among subgroups) KQ 4: Any active intervention listed above or no comparator (e.g., postmarketing surveillance study of an active intervention with no comparison group)<br>For double-blinded, placebo-controlled trials for network meta-analysis (all KQs): placebo | All other comparisons, including active interventions not listed above; dose-ranging studies that are not comparing two different interventions
Outcomes<br>KQs 1, 4: Disease activity, response, remission, radiographic joint damage<br>KQs 2, 4: Functional capacity, quality of life, patient-reported outcomes<br>KQs 3, 4: Overall risk of harms, overall discontinuation, discontinuation because of adverse effects, risk of serious adverse effects, specific adverse effects, patient adherence | All other outcomes not listed
Timing<br>All KQs: At least 3 months of treatment | <3 months treatment
Settings<br>All KQs: Primary, secondary, and tertiary care centers treating outpatients | Facilities treating inpatients only
Country setting<br>All KQs: Any geographic area | None
Study designs<br>For all KQs—i.e., benefits and harms overall (KQs 1, 2, 3) and among subgroups (KQ 4), study designs include head-to-head RCTs and nRCTs; prospective, controlled cohort studies (N>100); double-blinded, placebo-controlled trials for network meta-analysis; and SRs only to identify additional references.<br>For studies of harms—i.e., overall (KQ 3) and among subgroups (KQ 4), study designs also included any other observational study (e.g., cohort, case-control, large case series, postmarketing surveillance) (N>100). | All other designs not listed
Publication language<br>All KQs: English | Languages other than English

---

cs = conventional synthetic; DMARDs = disease-modifying antirheumatic drugs; KQ = Key Question; N = number; nRCT = nonrandomized controlled trial; PICOTS = population, intervention/exposure, comparator, outcomes, time frames, country settings, study design; RA = rheumatoid arthritis; RCT = randomized controlled trial; SR = systematic review; TNF = tumor necrosis factor; ts = targeted synthetic.

Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies To Answer the Key Questions

We systematically searched, reviewed, and analyzed the scientific evidence for each KQ. We included any study population defined as early RA by the authors if the diagnosis was no more than 1 year in the past. We included studies with mixed populations if more than 50 percent of the study populations had an early RA diagnosis. This definition was based on the context that the course of RA is highly variable; some researchers have suggested defining early RA as before development of bone erosion, but some patients never develop erosions. Given this variability, a recent task force of experts in RA and clinical trial methodology recommended defining early RA as no more than 1 year of diagnosed disease duration.²

Because no consensus on the definition of early RA exists, we also internally tracked studies with participants whose RA was between 1 to 2 years of diagnosis to describe the number of
studies using this time frame. If studies did not clearly indicate how early RA was defined but met our other PICOTS criteria, we attempted to contact the corresponding author to request clarification of the definition (using a standard email request). We gave authors 2 weeks to respond; if we did not receive a response after a reminder, we did not include the studies in question.

A portion of our literature yield consisted of abstract-only references without full-text manuscripts (e.g., conference abstracts). If we could not locate associated full-text publications, we excluded them because of a lack of information needed to assess risk of bias (ROB).

To identify relevant published literature, we searched the following databases: MEDLINE® via PubMed, the Cochrane Library, Embase, and International Pharmaceutical Abstracts. The search strategies formatted for MEDLINE (Appendix A) comprise medical subject heading (MeSH) terms and natural language terms reflective of RA, drug interventions, and outcomes of interest. We adapted this search strategy for the other databases as needed. An experienced librarian familiar with SRs designed and conducted all searches in consultation with the review team.

The 2012 review had searched from June 2006 to January 2011. For the present update, our literature searches included articles published from July 2010 (to allow 1 year’s indexing time from the 2012 update) to October 5, 2017.

We manually searched the reference lists of included SRs to supplement the main database searches. At the outset, we ensured that our update adequately builds on the body of evidence of the 2012 update, including new drugs approved by the U.S. Food and Drug Administration (FDA) or undergoing FDA review during our review period.

Because the scope of this update is limited to patients with early RA, we carefully examined included studies in the prior review to identify those that focused exclusively on patients with early RA or that had mixed populations of patients in which 50 percent or more had a diagnosis of early RA.

We also searched the gray literature for unpublished studies relevant to this review. Gray literature sources included ClinicalTrials.gov, the World Health Organization International Clinical Trials Registry Platform, the New York Academy of Medicine’s Grey Literature Index, and Supplemental Evidence and Data information from targeted requests and from a Federal Register Notice (public invitation posted in the Federal Register to submit relevant study data to AHRQ on behalf of Evidence-based Practice Centers [EPCs]). From these, we included studies that met all the inclusion criteria and contained enough methodological information to assess ROB. When we updated our published literature search, we also updated the gray literature searches.

To answer the Contextual Questions, we identified relevant literature opportunistically from our literature searches for KQs and used targeted literature searches to address remaining gaps in information.

Literature Review, Data Abstraction, and Data Management

To ensure accuracy, two reviewers independently reviewed all titles and abstracts. We used Abstrackr, an online citation screening tool, to review title and abstract records and manage the results. We then retrieved the full text for all citations deemed potentially appropriate for inclusion by at least one of the reviewers. Two team members independently reviewed each full-text article for eligibility. We resolved discrepancies by consensus or by involving a third, senior reviewer.
All results at both title/abstract and full-text review stages were tracked in an EndNote® bibliographic database (Thomson Reuters, New York, NY). Appendix B presents the list of studies excluded (with reasons) at the full-text level.

We designed, pilot-tested, and used a structured data abstraction form to ensure consistency of data abstraction. We abstracted data into categories that included (but were not limited to) the following: study design, eligibility criteria, intervention (drugs, dose, duration), additional medications allowed, methods of outcome assessment, population characteristics, sample size, attrition (overall and attributed to adverse events), results, and adverse event incidence. A second team member verified abstracted study data for accuracy and completeness.

Because studies often use more than one instrument to assess the same outcome, we established a hierarchy of outcome measures. We used this hierarchy to prioritize the information we abstracted. Table 3 documents this “priority” approach; preferred outcome measures are shown in bold. If study authors provided data for the preferred outcome measure, we did not abstract data from any other measure that assessed the same outcome. If no specific outcome measures are shown in bold in Table 3 within a category, we did not establish a hierarchy for that outcome.

Table 3. Outcomes and hierarchy of preferred measures for data abstraction

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Outcome Measures (Preferred Measures in Bold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQs 1,4 Disease activity</td>
<td>• DAS, DAS28, DAS-CRP (Disease Activity Score)</td>
</tr>
<tr>
<td></td>
<td>• SDAI</td>
</tr>
<tr>
<td></td>
<td>• Others</td>
</tr>
<tr>
<td>KQs 1,4 Response</td>
<td>• ACR 20/50/70 (American College of Rheumatology percentage improvement from baseline)</td>
</tr>
<tr>
<td></td>
<td>• EULAR response (based on DAS28 scores)</td>
</tr>
<tr>
<td></td>
<td>• Others</td>
</tr>
<tr>
<td>KQs 1,4 Remission</td>
<td>• Remission as defined by study (usually DAS28&lt;2.6 or DAS&lt;1.6 in prior report)</td>
</tr>
<tr>
<td>KQs 1,4 Radiographic joint damagea</td>
<td>• SHS (Sharp-van der Heijde method for scoring radiographic change)</td>
</tr>
<tr>
<td></td>
<td>• Larsen score change (radiographic measure)</td>
</tr>
<tr>
<td></td>
<td>• Others</td>
</tr>
<tr>
<td>KQs 2,4 Functional capacity</td>
<td>• HAQ, HAQ-DI-Health assessment questionnaire</td>
</tr>
<tr>
<td></td>
<td>• SOFI index</td>
</tr>
<tr>
<td></td>
<td>• Others</td>
</tr>
<tr>
<td>KQs 2,4 Quality of life</td>
<td>• SF-36</td>
</tr>
<tr>
<td></td>
<td>• EuroQoL EQ5D quality-of-life questionnaire</td>
</tr>
<tr>
<td></td>
<td>• Others</td>
</tr>
<tr>
<td>KQs 2,4 Patient-reported symptoms</td>
<td>• Any patient-reported symptoms</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Outcome Measures (Preferred outcome measures are in bold)</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------</td>
</tr>
</tbody>
</table>
| KQs 3,4 Overall risk of harms, overall discontinuation because of AEs, risk of serious AEs, specific AE, patient adherence | • Overall risk of harms  
• Overall discontinuation/discontinuation because of AEs/toxicity  
• Patient adherence  
• Risk of serious AEs (using FDA definition\(^{67}\))  
  - Life threatening  
  - Requires hospitalization  
  - Leads to lasting disability/congenital anomaly  
  - Or jeopardizes the patient in any other serious way  
• Specific AEs: Our focus was on the 11 events reported as most commonly occurring across all our eligible drugs according to their FDA-approved labels (organized in descending order from most to least common)  
  - Rash  
  - Upper respiratory tract infection  
  - Nausea  
  - Pruritus  
  - Headache  
  - Diarrhea  
  - Dizziness  
  - Abdominal pain  
  - Bronchitis  
  - Leukopenia  
  - Injection site reactions |

\(^{a}\) If studies reported progression based on MRI, we noted that in the Evidence Tables (Appendix C).

ACR 20/50/70 = American College of Rheumatology 20/50/70% improvement; AE = adverse event; DAS = Disease Activity Score based on 44 joints; DAS28 = Disease Activity Score based on 28 joints; DAS-CRP = Disease Activity Score based on C-Reactive Protein; EuroQoL EQ5D = European Quality of Life-5 Dimensions; EULAR = European League against Rheumatism; FDA = U.S. Food and Drug Administration; HAQ = Health Assessment Questionnaire (DI = Disability Index); KQ = Key Question; MRI = magnetic resonance imaging; SDAI = Simple Disease Activity Index; SF-36 = Medical Outcomes Study Short Form 36 Health Survey; SHS = Sharp/van der Heijde Method for Scoring Radiographs; SOFI = Signals of Functional Impairment Scale.

For adverse events, we abstracted data on overall adverse events, overall study discontinuation, discontinuation because of adverse events or toxicity, patient adherence, and any serious adverse events as defined by FDA.\(^{67}\) For head-to-head trials only, we abstracted data for the 11 specific adverse events (listed in Table 3) that are most commonly reported across all of our eligible drugs according to their FDA-approved labels.

**Assessment of Methodological Risk of Bias of Individual Studies**

To assess the ROB (i.e., internal validity) of studies, we used the Risk of Bias In Non-randomised Studies of Interventions (ROBINS-I)\(^{68}\) for nonrandomized controlled (nRCT) studies. We adapted the Cochrane ROB tool\(^{69}\) for randomized controlled trials (RCTs) by adding an item about the adequacy of intention-to-treat analyses of RCTs. We used predefined criteria based on the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.\(^{70}\) These included questions to assess selection bias, confounding, performance bias, detection bias, and attrition bias; concepts covered include adequacy of randomization, similarity of groups at baseline, masking, attrition, whether intention-to-treat analysis was used, method of handling dropouts and missing data, validity and reliability of outcome measures, and outcome reporting bias.\(^{63}\) To assess outcome reporting bias, we checked protocols for eligible studies in
ClinicalTrials.gov ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)) when available, to determine which outcomes of a specific study had been registered a priori.

Two independent reviewers assessed ROB for each study. Disagreements between the two reviewers were resolved by discussion and consensus or by consulting a third member of the team.

**Data Synthesis**

We summarized all included studies in narrative form and in summary tables that tabulate the important features of the study populations, design, intervention, outcomes, setting, country, geographic location, and results. All new qualitative and quantitative analyses synthesized included relevant studies from the 2012 SR.

We considered performing pairwise meta-analyses for outcomes with information from at least three unique studies of low or medium ROB that we deemed to be sufficiently similar (in population, interventions, comparators, and outcomes). However, because of a lack of similar head-to-head trials, we were unable to conduct pairwise meta-analyses for any of the comparisons of interest. To address the dearth of studies directly comparing interventions of interest, we considered network meta-analyses. We assessed patient and study characteristics across studies that compared pairs of treatments to ensure the transitivity assumption (i.e., that potential effect modifiers are similar across studies) would hold. To be eligible for network meta-analyses, included studies had to fulfill the following four criteria: (1) patients with early RA had not attempted prior treatment with MTX; (2) doses of treatments were within FDA-approved ranges; (3) length of followup was similar; and (4) studies were double-blinded RCTs of low or medium ROB. Head-to-head and placebo-controlled RCTs were eligible for network meta-analyses; however, we did not find any eligible placebo-controlled trials in a population with early RA. We considered network meta-analyses for the following outcomes: American College of Rheumatology 50% improvement (ACR50), Disease Activity Score (DAS) remission, radiographic joint damage, all discontinuations from the study, and discontinuations attributed to adverse events.

Studies that we had rated high ROB were excluded from these analyses; we used them only in sensitivity analyses. We describe their findings briefly in the context of our main analyses.

We collected data on the number of participants and the number of events for each treatment group for dichotomous outcomes (ACR50, DAS, and discontinuations). For our sole continuous outcome analyzed (radiographic joint damage), we collected means and standard deviations (SDs) from the pre- and post-treatment time point for each study. Four studies did not have data for post-treatment SDs for radiographic joint damage; therefore, we imputed these data by pooling post-treatment SDs from four other studies. SDs for MTX were imputed by pooling SDs from the MTX arms of those studies (N=963 patients), while SDs for the other treatments were imputed by pooling SDs for the other treatment arms of those studies (N=1,730 patients).

We ran our network meta-analyses using a multivariate random effects meta-regression model with restricted maximum likelihood estimation. Models were fit using the Network package in Stata ([StataCorp, College Station, TX](http://www.stata.com)). This approach accounts for multiarm trials. We provide diagrams outlining the structure of the network for each outcome, with the lines in the diagrams representing direct comparisons between treatments and the size of the nodes for each treatment being proportional to the sample size. For closed loops, we tested the transitivity assumption by comparing consistency and inconsistency models and network side splits.
Because the global Wald test indicated significant differences between the consistency and inconsistency models, we presented the estimates from the consistency model.

We summarize results for dichotomous outcomes (ACR50, DAS, and discontinuations) in forest plots using relative risks. For the sole continuous outcome analyzed (radiographic joint damage), we report standardized mean differences (mean difference divided by standard deviation). We did not calculate ranking probabilities for treatments because such rankings may exaggerate small differences in relative effects.

We also carefully explored whether treatment strategies used for average patients with early RA can be used effectively or safely for patients with significant coexisting ailments such as hepatitis C, congestive heart failure, cancer, diabetes, and others. Because we lacked access to individual patient data, we used a qualitative approach to address this question.

**Grading the Strength of Evidence for Major Comparisons and Outcomes**

We graded the strength of evidence (SOE) based on the guidance established for the EPC Program. Developed to grade the overall strength of a body of evidence, this approach incorporates five key domains: (1) study limitations (including study design and aggregate ROB), (2) consistency, (3) directness, (4) precision of the evidence, and (5) reporting bias. It also considers other optional domains that may be relevant for some scenarios. These included plausible confounding that would decrease the observed effect and strength of association (i.e., magnitude of effect) or factors that would increase the strength of association (i.e., dose-response effect). To grade the SOE of results from network meta-analysis, we used guidance from the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group. The SOE for indirect estimates was downgraded for indirectness and imprecision in all cases. For comparisons that had both direct and indirect evidence, we commented on whether the indirect evidence was consistent with the direct evidence.

Table 4 describes the grades of evidence that can be assigned. Grades reflect the strength of the body of evidence to answer the KQs on the comparative effectiveness, efficacy, and harms of the interventions in this review. Two reviewers assessed each domain for each key outcome, and they resolved differences by consensus discussion.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).</td>
</tr>
<tr>
<td>Moderate</td>
<td>We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.</td>
</tr>
<tr>
<td>Low</td>
<td>We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.</td>
</tr>
<tr>
<td>Insufficient</td>
<td>We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.</td>
</tr>
</tbody>
</table>

Source: Berkman et al., 2014.
We graded the SOE for the following outcomes, consistent with the prior report: disease activity, response, radiographic joint damage, functional capacity, discontinuation because of adverse events, and serious adverse events.1

Assessing Applicability

We assessed the applicability of individual studies and the larger body of evidence, following guidance from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews.75 We examined the following points: whether interventions were similar to those in routine use, whether comparators reflected best alternatives, whether measured outcomes reflected the most important clinical outcomes, whether followup was sufficient, and whether study settings were representative of most outpatient settings. For individual studies, we examined conditions that may limit applicability based on the PICOTS structure. In particular, we focused on factors such as race or ethnicity of populations in studies, clinical setting, geographic setting, and availability of health insurance.

Peer Review and Public Commentary

The AHRQ Task Order Officer and an AHRQ associate editor (a senior member of another EPC) reviewed the draft report before peer review and public comment. The draft report (revised as needed) was sent to invited peer reviewers and simultaneously uploaded to the AHRQ Web site where it was available for public comment for 52 days with a 1-week holiday-related extension.
Results

Organization of the Results

We first present the results of the literature search and provide a literature flow diagram. In the Characteristics of Included Studies section, we report the distribution of studies by study design and drug therapy group across the Key Questions (KQs). Because most of the included studies provide results for multiple KQs, we describe the study and participant characteristics only once before reporting the KQ-specific results. These characteristics are organized by drug therapy group and drug therapy comparison subgroups. Then, we provide KQ-specific results, which are organized in the same manner. To recap, KQ 1 and KQ 2 deal with benefits of therapy, measured by intermediate or final health outcomes, respectively; KQ 3 focuses on harms of therapy; and KQ 4 addresses issues relating to subpopulations.

Evidence tables that include additional details on study and population characteristics and outcomes appear in Appendix C, followed by study risk of bias (ROB) assessments in Appendix D, outcome-level strength of evidence (SOE) grading details in Appendix E, a description of eligible clinical assessment scales used in our body of evidence and their scoring in Appendix F, detailed test of consistency results for our primary network meta-analyses (NWMA) in Appendix G, the results of supplementary primary NWMA not presented in the main report in Appendix H, and the results of our sensitivity analyses for NWMA in Appendix I.

Search Results

Our electronic searches identified 6,373 citations (Figure 2). We identified an additional 429 citations through other sources; these included the prior report, team member or reviewer recommendations, handsearching of relevant systematic reviews, companion article additions, and supplemental evidence and data received through the Agency for Healthcare Research and Quality (AHRQ) Web site and a Federal Register notice. Following initial removal of duplicate records (details available in Appendix A), a total of 5,287 unique citations underwent title and abstract screening. Of those, 1,628 required full-text review, and 49 studies reported in 124 articles (3% total yield) met our eligibility criteria for inclusion in this review.

Characteristics of Included Studies

In total, 49 studies reported in 124 articles were included; we had 41 randomized controlled trials (RCTs), 4 comparative observational studies, and 4 single-arm observational studies. We grouped studies by the main drug therapy group being evaluated: corticosteroids, conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), targeted synthetic DMARDs (tsDMARDs), tumor necrosis factor (TNF) biologics, non-TNF biologics, biosimilars, and combinations and therapy strategies.

We use Tables 5 through 10 to describe our evidence base and present individual study results. Table 5 presents the distribution of studies by study design and drug therapy group across the KQs. Table 6 presents an overview of important details about our review’s evidence base.
Tables 7, 9, and 10 report major findings from studies used to answer KQ 1, KQ 2, and KQ 3, respectively. Tables 8 and 11 provide a summary of details for all studies used in our KQ 1 and KQ 3 NWMA, including their treatment comparisons and specific outcomes for which they were analyzed. Appendix C provides additional study and population characteristics and outcomes.

Within each drug therapy group, we further categorized studies based on the comparisons that any given study was evaluating (e.g., a csDMARD monotherapy versus a different csDMARD monotherapy). Below, we describe study and patient characteristics for the included studies, grouped by the main drug therapy and then by the comparison(s) the authors made. Patient characteristics were similar by randomized groups; studies with any baseline differences were rated as having a higher risk of bias.

**Figure 2. Summary of literature search flow and yield for early rheumatoid arthritis**
IPA = International Pharmaceutical Abstracts; NWMA = network meta-analysis; NY = New York; RA = rheumatoid arthritis; SEADs = supplemental evidence and data; WHO ICTRP = World Health Organization International Clinical Trials Registry Platform; yr = year.
<table>
<thead>
<tr>
<th>Drug Therapy Group</th>
<th>Comparison Type</th>
<th>Overall N of Studies</th>
<th>KQ 1 Intermediate Outcomes</th>
<th>KQ 2 Final Health Outcomes</th>
<th>KQ 3 Harms</th>
<th>KQ 4 Subpopulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroids</td>
<td>Corticosteroids vs. csDMARDs</td>
<td>6</td>
<td>6 RCTs</td>
<td>5 RCTs</td>
<td>6 RCTs</td>
<td>None</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>High-dose corticosteroid vs. TNF biologic</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2 RCTs</td>
<td>2 RCTs</td>
<td>2 RCTs</td>
<td>None</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>Corticosteroid single-arm studies</td>
<td>1</td>
<td>None</td>
<td>None</td>
<td>1 obs</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARD monotherapy vs. csDMARD monotherapy</td>
<td>2</td>
<td>2 studies (1 RCT, 1 obs)</td>
<td>2 studies (1 RCT, 1 obs)</td>
<td>2 studies (1 RCT, 1 obs)</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARD combination therapy vs. csDMARD monotherapy</td>
<td>7</td>
<td>7 studies (6 RCTs, 1 obs)</td>
<td>6 RCTs</td>
<td>7 studies (6 RCTs, 1 obs)</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARDs vs. TNF biologics</td>
<td>1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARDs vs. non-TNF biologics</td>
<td>3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3 RCTs</td>
<td>3 RCTs</td>
<td>3 RCTs</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARDs vs. tsDMARDs</td>
<td>1</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>None</td>
</tr>
<tr>
<td>csDMARDs</td>
<td>csDMARD single-arm studies</td>
<td>4</td>
<td>None</td>
<td>None</td>
<td>4 obs</td>
<td>None</td>
</tr>
<tr>
<td>Biologics</td>
<td>Biologics vs. csDMARD monotherapies</td>
<td>16&lt;sup&gt;a,b,c&lt;/sup&gt;</td>
<td>16 RCTs</td>
<td>16 RCTs</td>
<td>15 RCTs</td>
<td>3 RCTs</td>
</tr>
<tr>
<td>Biologics</td>
<td>Biologics vs. csDMARD combination therapies</td>
<td>3</td>
<td>3 RCTs</td>
<td>3 RCTs</td>
<td>3 RCTs</td>
<td>1 RCT</td>
</tr>
<tr>
<td>Biologics</td>
<td>Biologic head-to-head comparisons</td>
<td>1</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>1 RCT</td>
<td>None</td>
</tr>
<tr>
<td>Biologics</td>
<td>Biologic single-arm studies</td>
<td>1</td>
<td>None</td>
<td>None</td>
<td>1 obs</td>
<td>None</td>
</tr>
<tr>
<td>Combination and therapy strategies</td>
<td>N/A</td>
<td>6</td>
<td>4 RCTs</td>
<td>4 RCTs</td>
<td>6 studies (4 RCTs, 2 obs)</td>
<td>None</td>
</tr>
</tbody>
</table>

<sup>a</sup> One study evaluated comparisons relevant to two categories: high-dose corticosteroid vs. TNF biologic and biologic vs. csDMARD monotherapies.<sup>18</sup>  
<sup>b</sup> One study evaluated comparisons relevant to two categories: csDMARD vs. TNF biologics and biologics vs. csDMARD monotherapies.<sup>15</sup>  
<sup>c</sup> Three studies evaluated comparisons relevant to two categories: csDMARD vs. non-TNF biologics and biologics vs. csDMARD monotherapies.<sup>7,32,33</sup>  

csDMARD = conventional synthetic disease-modifying antirheumatic drug; KQ = Key Question; N = number; N/A = not applicable; obs = observational study(ies); RCT = randomized controlled trial; TNF = tumor necrosis factor; tsDMARD = targeted synthetic disease-modifying antirheumatic drug; vs. = versus.
Table 6. Characteristics of included studies

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Corticosteroids</th>
<th>csDMARDs and tsDMARDs</th>
<th>Biologics</th>
<th>Any Combinations and Therapy Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of studies (articles)</td>
<td>9 (14)</td>
<td>18 (40)</td>
<td>23 (62)</td>
<td>6 (23)</td>
</tr>
<tr>
<td>N of studies (articles) included in prior report</td>
<td>2 (3)</td>
<td>6 (12)</td>
<td>6 (14)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Countries</td>
<td>Belgium, England/Wales, Germany, Italy, Netherlands, Sweden, United Kingdom</td>
<td>Australia, Belgium, Denmark, Finland, France, Germany, multinational (not specified), Netherlands, Norway, Sweden</td>
<td>Argentina, Austria, Belgium, Canada, Colombia, Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Japan, multinational (not specified), Mexico, Monaco, Netherlands, Poland, Romania, Spain, Sweden, Switzerland, United Kingdom, United States</td>
<td>Denmark, France, Ireland, Netherlands, United Kingdom, United States</td>
</tr>
<tr>
<td>N of patients</td>
<td>14,586</td>
<td>37,536</td>
<td>22,590</td>
<td>4,375</td>
</tr>
<tr>
<td>Sex: range of % female patients</td>
<td>60 to 80.9</td>
<td>58.3 to 82.6</td>
<td>53.4 to 81.4</td>
<td>65 to 80</td>
</tr>
<tr>
<td>Age: range of means</td>
<td>50 to 62</td>
<td>47 to 63.8</td>
<td>46 to 58</td>
<td>46.3 to 58</td>
</tr>
<tr>
<td>Disease duration: N (%) enrolling only patients with early RA (≤1 year)</td>
<td>9 (100)</td>
<td>9 (50)</td>
<td>11 (47.8)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>Study durations</td>
<td>1 to 15 years</td>
<td>6 months to 15 years</td>
<td>6 months to 15 years</td>
<td>1 to 10 years</td>
</tr>
<tr>
<td>ROB (N of studies)</td>
<td>Medium: 8&lt;sup&gt;b,c&lt;/sup&gt; High: 1&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>Medium: 11&lt;sup&gt;b&lt;/sup&gt; High: 3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Low: 7&lt;sup&gt;d&lt;/sup&gt; Medium: 14&lt;sup&gt;e&lt;/sup&gt; High: 7&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Low: 1&lt;sup&gt;f&lt;/sup&gt; Medium: 3&lt;sup&gt;g&lt;/sup&gt; High: 4&lt;sup&gt;h&lt;/sup&gt;</td>
</tr>
<tr>
<td>N of studies (articles) reporting on benefits (KQ 1 or 2)</td>
<td>8 (13)</td>
<td>14 (31)</td>
<td>22 (61)</td>
<td>4 (22)</td>
</tr>
<tr>
<td>N of studies (articles) reporting on harms (KQ 3)</td>
<td>8 (13)</td>
<td>18 (35)</td>
<td>22 (60)</td>
<td>6 (24)</td>
</tr>
<tr>
<td>N of studies (articles) reporting on subgroup effects (KQ 4)</td>
<td>0</td>
<td>0</td>
<td>4 (17)</td>
<td>0</td>
</tr>
<tr>
<td>Specific drugs evaluated (N of studies for each)</td>
<td>Methyl-PNL: 2; PRED: 4; PNL: 2; Oral CS (not specified): 1</td>
<td>LEF: 1; MTX: 14; SSZ: 7; TOF: 1; HCQ+MTX+SSZ: 2; csDMARD combo (not specified): 1</td>
<td>ABA: 2; ADA: 5; CZP: 2; ETN: 3; IFX: 5; RIT: 2; TCZ: 2; TNF biologics (not specified): 1</td>
<td>N/A (see Table 7, Table 9, or Table 10 for specific drug combinations)</td>
</tr>
</tbody>
</table>
Also within each drug therapy group, we provide the number of studies enrolling samples made up entirely of early RA patients with a disease duration ≤1 year, as well as the number of studies that enrolled mixed populations of patients with early RA.

The range of mean or median disease durations across all 49 included studies was 2 weeks to 12 months. Prior treatment use varied widely across drug therapy categories. Among all 49 included studies, five studies did not report any details about prior treatment use, leaving 44 studies that did. Of these, 36 enrolled MTX-naïve patient samples, and the remaining eight studies enrolled patients with at least some prior csDMARD use (including MTX).

In four of these eight studies, prior use of any csDMARDs ranged from 13 to 48 percent. The other four enrolled samples that were entirely csDMARD resistant. Among the 15 studies analyzed in our primary or sensitivity NWMA, five enrolled patients with some prior

A

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Corticosteroids</th>
<th>csDMARDs and tsDMARDs</th>
<th>Biologics</th>
<th>Any Combinations and Therapy Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drugs not evaluated</td>
<td>None</td>
<td>None</td>
<td>ADA-atto, ETN-szsz, GOL, IFX-abda, IFX-dyyb, SAR</td>
<td>N/A</td>
</tr>
</tbody>
</table>

.a Study counts in the corticosteroids, csDMARDs, and/or biologic categories share several studies that have evaluated comparisons pertaining to multiple drug categories. We did not assign ROB ratings to single-arm studies reporting on harms, including one study of corticosteroids and four studies of csDMARDs. One study of corticosteroids had two ROB ratings for outcomes at different time points. We assigned a medium rating to 1-, 2-, and 10-year outcomes and a high rating to 4-year outcomes. One study of biologics (AGREE) received both low and medium ROB ratings that were outcome specific. We assigned a low rating to ACR response, DAS28 remission, LDAS, radiographic outcomes, and AEs and a medium rating was assigned to HAQ-DI and SF-36. Five studies of biologics received both medium and high ROB ratings that were specific to either outcomes or time points. Among the two studies with outcome-specific ratings, we assigned a medium rating to DAS28 remission, ACR response, HAQ-DI, and SF-36 and a high rating to mTSS and SHS erosion scores in one study (HIT HARD). We assigned a medium rating to all outcomes, except for WPS-RA work productivity outcomes, which were reported only on ClinicalTrials.gov and received a high rating. Among the three studies with time point-specific ratings, we assigned medium ratings to 16-week outcomes in one study (PROWD), 24-week outcomes in a second (C-OPERA), and 52-week outcomes in the third. We assigned high ratings to 52-week outcomes in one study (PROWD) and 52-week outcomes in the other two (C-OPERA and FUNCTION). One study of combination and therapy strategies (BeSt) received both low and medium ROB ratings that were time point-specific. Among the four studies of combination and therapy strategies receiving high ROB ratings, only one had different ratings for specific time points. We assigned this study (GUEPARD) a high rating only for its 52-week outcomes, but a medium rating for its 12-week outcomes. One head-to-head study of TNF biologics vs. non-TNF biologics evaluated both RIT and ADA or ETN, but without isolating the effects of either drug.

ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; AE = adverse event; combo = combination; csDMARD = conventional synthetic disease-modifying antirheumatic drug; CZP = certolizumab pegol; DAS28 = Disease Activity Score based on 28 joints; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; GOL = golimumab; HAQ = Health Assessment Questionnaire (DI = Disability Index); HCO = hydroxychloroquine; IFX = infliximab; KQ = Key Question; LDAS = low disease activity score; LTF = leflunomide; Methyl-PNL = methylprednisolone; mTSS = modified Total Sharp/van der Heijde score; MTX = methotrexate; N = number; N/A = not applicable; PNL = prednisolone; PRED = prednisone; RA = rheumatoid arthritis; RIT = rituximab; ROB = risk of bias; SAR = sarilumab; SF-36 = Short-Form Health Survey 36-Item; SHS = Sharp/van der Heijde Score; SSZ = sulfasalazine; TCZ = tocilizumab; TOF = tofacitinib; tsDMARD = targeted synthetic disease-modifying antirheumatic drug; WPS-RA = Work Productivity Survey - Rheumatoid Arthritis.
csDMARD use other than MTX,12-16 and three did not report whether patients had used other csDMARDs.7, 17, 18

Five of the eight studies enrolled samples that had previously used MTX specifically: 5819 and 7920 percent of patients in two studies, and three studies (all trials) enrolling samples that were entirely MTX resistant (i.e., 100% prior use).8-10

All included studies enrolled patients with moderate to high disease activity at baseline as measured with mean or median Disease Activity Score (DAS) 28 scores (range of 0 to 10). More than one-half (53% to 83%) of the patient population was women. The mean age range was 46 to 64 years. Study durations ranged from 6 months to 15 years.

Corticosteroids

We included eight RCTs3, 6, 18, 78, 93-96 and one single-arm observational study76 that evaluated corticosteroids. Of the eight RCTs, all contributed results to KQs 1 and 3, and six contributed results to KQ 2. The one single-arm observational study contributed only to KQ 3. Two corticosteroid studies (three articles)78, 93, 97 had been included in the prior report1 (Table 6).

All nine corticosteroid studies enrolled samples consisting entirely of early RA patients with disease duration ≤1 year.3, 6, 18, 76, 78, 93-96

Corticosteroids Versus csDMARDs

Six RCTs compared corticosteroids with csDMARDs (Appendix C). Each took place in various European countries over 2 years (except for one3 that lasted only 1 year). Four trials compared a combination of prednisone (PRED) and MTX versus MTX alone.3, 6, 94, 95, 98, 99 One of these four trials evaluated this comparison in patients at low risk of poor disease prognosis; patients in this trial at high risk of a poor prognosis received additional treatment with either sulfasalazine (SSZ) or leflunomide (LEF) on top of combination PRED and MTX.95, 98 As for the remaining two trials, one evaluated a combination of prednisolone (PNL) and MTX versus MTX alone;93 the other compared a combination of PNL and a csDMARD (mostly MTX or SSZ) versus csDMARD monotherapy.78

Most of the patients in these RCTs were female (60% to 81%), with a mean age between 51 and 62 years. Their disease durations were generally similar and ranged from a mean or median of 2.7 to 6.5 months; one study’s patients had a notably shorter mean duration of less than a month (1.8 to 3.2 weeks).95

Mean baseline DAS values ranged from 3.7 to 5.9, and mean baseline Health Assessment Questionnaire (HAQ) ranged from 1.0 to 1.7. Four studies reported mean baseline Sharp scores: three reported similar mean or median scores ranging from 0.7 to 1.3, but the fourth had notably higher mean scores (4.1 to 4.8) (see Appendix F for a description of scales).

Four studies reported information about prior use of MTX or other csDMARDs.78, 93-95 In the three studies reporting on MTX use, all patients were MTX naïve.78, 94, 95 Among the four studies reporting on prior csDMARD use, three recruited patients who were csDMARD naïve,78, 94, 95 and a small proportion of patients (about 14 percent) in one study had a history of DMARD use.93

High-Dose Corticosteroids

Two RCTs from Belgium and the United Kingdom (lasting 52 to 78 weeks) compared a combination of a high-dose corticosteroid, namely IV methyl-PNL (doses of 250 mg96 or 1 g18),
and MTX versus a combination of infliximab (IFX) and MTX. Additionally, one study compared the combination of high-dose methyl-PNL and MTX versus MTX monotherapy. Moreover, one study compared the combination of high-dose methyl-PNL and MTX versus MTX monotherapy.18

Most of the trials’ patients were female (67% and 71%, respectively); the mean age of all patients across treatment arms ranged from 50 to 54 years. The disease duration was a median of 1.2 months in the United Kingdom study, and a mean of nearly 6 months. across treatment arms. Mean baseline DAS ranged from 3.6 to 5.3 across treatment arms. Across studies, mean baseline HAQ ranged from 1.3 to 1.5, and the average baseline Sharp score was only reported in one study, ranging from 6.1 to 9.2 across treatment arms. Both studies’ patients were entirely MTX naïve, and one reported on csDMARD use in general, specifically that its sample was csDMARD naïve.

Corticosteroids: Single-Arm Studies

One study from Sweden (lasting 15 years) evaluated harms associated with oral corticosteroids used for patients with early RA. The range of oral corticosteroid doses used by patients was not measured over the course of the study, but rather, only their use or non-use during the first year after RA diagnosis.

Most of the study’s patients were female (69%), with a mean age of 58 years. The mean disease duration was not reported, but all patients’ disease durations were less than 1 year. Median baseline DAS was 5.2, but neither mean baseline HAQ nor Sharp scores were reported. All study participants had no history of prior treatment with MTX or csDMARDs in general.

csDMARD Studies

We included 11 RCTs, comparative observational studies, and 4 single-arm observational studies that evaluated csDMARDs. All 11 RCTs contributed results to KQs 1, 2, and 3. Overall, we used five of these RCTs in our NWMA. Both comparative observational studies contributed to KQs 1 and 3, but only one contributed to KQ 2. Each single-arm observational study contributed only to KQ 3. Six csDMARD studies (12 articles) had also been included in the prior report (Table 6). Most of our csDMARD studies (n=8) enrolled mixed populations in terms of RA disease duration. The remaining nine enrolled samples were made up entirely of early RA patients with disease duration ≤1 year.

csDMARDs Versus csDMARDs

Seven RCTs and two single-arm observational studies compared csDMARD monotherapies versus either other csDMARD monotherapies or csDMARD combination therapies. Appendix C describes all these studies in detail. The studies took place mainly in European countries; five were based in the Netherlands. Intervention details and characteristics are summarized below by type of csDMARD drug (e.g., monotherapy or combination).

csDMARD Monotherapy Versus csDMARD Monotherapy

One RCT and one prospective cohort study compared csDMARD monotherapies versus other csDMARD monotherapies. Each took place over 2 to 3 years in Sweden or Norway. The RCT compared the efficacy of two different csDMARDs, MTX versus SSZ, both combined with PRED. The cohort study evaluated the same comparison (MTX versus SSZ), but did not use PRED in combination.
The patients in both studies were similar in terms of demographics: mean ages were approximately 50 and 54 years, and most patients were female (63% and 67%). Only the RCT reported disease duration at baseline, a median of 6 months. Mean baseline DAS was 4.4 and 5.0, and median baseline HAQ 0.5 and 0.9; neither study reported mean Sharp score. In terms of prior treatment history, all patients in both the RCT and observational study were MTX and csDMARD naïve.

**csDMARD Combination Therapy Versus csDMARD Monotherapy**

We included six RCTs\(^4,21-24,105\) and one prospective cohort study\(^26\) comparing csDMARD monotherapies versus csDMARD combination therapies. Each took place over 1 to 5 years across multiple countries. The RCTs compared the efficacy of multiple csDMARDs combined with each other (plus PNL or other glucocorticoids in four studies\(^4,22,24,105\)) versus MTX or SSZ monotherapy.\(^4,21-24,105\) The cohort study compared the combination of SSZ and MTX versus MTX alone.\(^26\)

Patients varied across studies in terms of demographics: mean ages in the RCTs ranged from 47 to 57 years; the cohort study’s sample had a mean age ranging from approximately 62 to 64 years across treatment arms. Most patients in each study were female (range of 58% to 77%). Disease duration at baseline varied from a mean of 2.3 months to a median of nearly a year (47 weeks). Mean baseline DAS was 3.6 to 5.7, and mean baseline HAQ ranged from 0.9 to 1.4. Four studies reported Sharp scores, which varied considerably across studies from a median of 0 to a mean of 8.9.\(^4,21,24,25\)

Prior treatment history was reported for MTX use in five of the RCTs\(^21-25\) and csDMARD use in three RCTs.\(^21-23\) Among these RCTs, all patients were treatment naïve. Only a small proportion of the prospective cohort’s sample reported prior use of csDMARDs (range of 13% to 15% across treatment arms).\(^26\)

**csDMARDs Versus Biologics**

Four RCTs compared csDMARD monotherapies versus biologics.\(^7,15,32,33\) Three trials were multinational;\(^7,15,32\) one was based solely in The Netherlands.\(^33\) Appendix C summarizes the intervention details and patient characteristics of these trials.

**csDMARDs Versus TNF Biologics**

One multinational RCT compared the combination of a csDMARD (MTX) and a TNF biologic (adalimumab [ADA]) versus ADA alone and MTX alone. The study took place over 2 years.\(^15\)

Patients enrolled in this trial had a mean age of approximately 52 years. Most of the sample was female (74.5%). As for prior treatment history, most patients were treatment-naïve, with the entire sample being MTX-naïve and about one-third reporting prior csDMARD use (32%).

**csDMARDs Versus Non-TNF Biologics**

We included three RCTs comparing csDMARDs with non-TNF biologic monotherapies or combined with csDMARDs. One RCT compared the combination of a csDMARD (MTX) and a non-TNF biologic (abatacept [ABA]) versus ABA alone and MTX alone.\(^7\) Two RCTs compared the combination of a csDMARD (MTX) and a non-TNF biologic (tocilizumab [TCZ]) versus TCZ alone and MTX alone. The trials took place over 1 to 2 years.\(^7,32,33\)
Patients in these three trials had mean ages of 47 and 54 years, and most patients were female (range of 67% to 78% across treatment arms). Median disease duration at baseline ranged from 1 to 6 months. Mean DAS scores at baseline were between 5.2 and 6.7 across treatment arms, and mean HAQ scores at baseline were 1.2 to 1.75. Mean Sharp scores varied notably between the only two studies reporting these baseline data, with a median of 0.0 in one \(^33\) and means ranging from 5.7 to 7.7 across the other study’s treatment arms.\(^32\)

Both samples were treatment-naïve in terms of previous MTX or other csDMARD use. Two studies targeted treatment of aggressive early RA.\(^7, 32\) In one, 89.5 percent of its sample was rheumatoid factor (RF) seropositive, and its entire sample was experiencing erosive disease.\(^32\) In the other trial, 72 percent of the sample was RF seropositive.\(^33\)

**csDMARDs Versus tsDMARDs**

One multinational RCT (lasting 1 year) compared the combination of tofacitinib (TOF) and MTX versus TOF alone and MTX alone (Appendix C).\(^29\)

Patients enrolled in this study had a mean age of approximately 48 to 51 years across treatment arms. Most of the sample was female (about 83%). Mean DAS scores ranged from 6.3 to 6.5 across treatment arms, and the overall mean HAQ score was 1.5. Mean Sharp scores ranged from 12.6 to 13.7 across treatment arms.

As for prior treatment history, very few reported prior MTX use (5.5%), and no information about previous csDMARD use in general was available.

**csDMARDs: Single-Arm Studies**

Four single-arm studies evaluated harms associated with csDMARDs (Appendix C).\(^5, 19, 76, 77\) Study duration varied widely: a mean of 25 weeks in one study,\(^19\) a median of 2 years in another,\(^5\) about 8 years in a third,\(^77\) and 15 years in a fourth.\(^76\) Three studies took place in European countries;\(^19, 76, 77\) the third was based in Australia.\(^5\)

Most of the studies’ patients were female (about 67% to 73%), with a mean age of approximately 53 to 60 years. The disease duration was reported by three of these studies\(^5, 19, 77\) and ranged from a median of 4 months to approximately 8 months; only one study\(^19\) reported a mean disease duration, which was 7.5 months. Information about prior treatment was reported in two studies;\(^19, 76\) in one,\(^19\) slightly less than one-half of the sample reported prior MTX or csDMARD use, and in the other,\(^76\) the entire sample was treatment naïve.

**Biologics**

We included 22 RCTs and one single-arm study that evaluated TNF and non-TNF biologics. All but one\(^76\) contributed results to KQs 1 and 2, all but one\(^41\) contributed results to KQ 3, and four reported eligible data for KQ 4\(^10, 14, 17, 35\) (Appendix C). Five biologic DMARD studies (12 articles)\(^10, 12, 14, 17, 31, 106-112\) had also been included in the prior report\(^1\) (Table 6).

Most of our trials of biologics (n=12) enrolled mixed populations of early RA patients and those with longer-duration RA.\(^7-9, 12, 14-17, 30-32, 35\) The remaining 10 studies enrolled samples made up entirely of early RA patients with disease duration ≤1 year.\(^10, 13, 18, 34, 37, 38, 40, 41, 76, 113\)

**TNF Biologics Versus csDMARDs**

We included 16 RCTs comparing TNF biologics versus csDMARDs (Appendix C).\(^9, 10, 12-18, 34, 35, 37, 38, 40, 41, 113\) Eight were conducted solely in European countries;\(^9, 10, 16, 18, 34, 40, 41, 113\) two were based in Japan\(^13, 35\) and one in the United States;\(^14\) five were multinational.\(^12, 15, 17, 37, 38\)
Intervention details and characteristics are summarized below by whether studies used csDMARD monotherapy or combination therapy as the comparator.

**TNF Biologic Versus csDMARD Monotherapy**

Thirteen RCTs compared TNF biologics versus csDMARD monotherapy.\(^{12-18, 34, 35, 37, 38, 41, 103, 113-119}\) Trials lasted from 6 months to 2 years. Five trials compared a combination of ADA and MTX versus MTX alone.\(^{15, 16, 34, 35, 37}\) One used an MTX dose lower than the dose currently approved by the U.S. Food and Drug Administration.\(^{35}\) Two trials compared etanercept (ETN) versus MTX alone,\(^{12, 113}\) and one evaluated the combination of ETN and MTX versus MTX alone.\(^{14}\) Another three compared a combination of IFX and MTX versus MTX alone.\(^{17, 18, 41}\) Two trials compared a combination of certolizumab pegol (CZP) and MTX versus MTX alone.\(^{13, 38}\) We included nine of these RCTs in our NWMA.

Patients in these trials were mostly female (53% to 81%) with a mean age between 47 years and 54 years. Their mean duration of disease was reported in all but one study and varied from about 2 months to 12 months. Baseline DAS ranged from a mean or median of 5.2 to 6.9, and mean baseline HAQ ranged from 1.0 to 1.9. Mean baseline Sharp scores ranged across studies from 2.4 to 22.

All 13 trials of TNF biologics enrolled samples of MTX-naïve patients, but the proportion of patients reporting other prior treatments differed across studies. Eleven trials reported information about prior treatment, specifically csDMARDs (as a broad category). Four trials enrolled samples of csDMARD-naïve patients,\(^{34, 38, 41, 113}\) five reported that approximately 18 to 54 percent of their patients had taken any csDMARDs,\(^{12-14, 35, 37}\) and one reported that its patients used a mean of 0.2 csDMARDs at baseline.\(^{16}\) The two trials not reporting prior csDMARD use did not differ in a notable way from the other TNF biologic studies.\(^{17, 18}\)

**TNF Biologic Versus csDMARD Combination Therapy**

Three RCTs compared TNF biologics versus csDMARD combination therapy.\(^{9, 10, 40, 120-128}\) Each trial lasted 2 years. All three trials compared a combination of TNF biologics and csDMARDs versus a three- or four-drug combination therapy; however, no trial evaluated the same exact combination. One trial compared a combination of MTX, PRED, hydroxychloroquine (HCQ), and SSZ versus MTX and ADA.\(^{9, 120}\) Another compared the combination of IFX and the FIN-RACo (Finnish Rheumatoid Arthritis Combination Therapy trial) regimen (MTX, PRED, HCQ, and SSZ) versus the FIN-RACo regimen alone.\(^{40, 127, 128}\) The third trial compared triple therapy of MTX, SSZ, and HCQ versus a combination of MTX and IFX.\(^{10, 121-126}\)

Patients in these RCTs were mostly female (67% to 79% across treatment arms), with a mean age between 46 and 53 years. Their mean disease durations ranged from approximately 4 to 6 months. Baseline DAS ranged from a mean of 2.5 to 5.6, and mean baseline HAQ ranged from 0.9 to 1.3.

Two trials enrolled patients who had all previously used MTX,\(^{9, 10}\) and patients in the third reported no prior treatment with MTX or csDMARDs.\(^{40}\)

**TNF Biologics: Single-Arm Studies**

One study from Sweden (lasting 15 years) evaluated harms associated with TNF biologics used for patients with early RA.\(^{76}\) This study has also been described previously in the
Corticosteroids and csDMARDs sections because it evaluated harms for drugs within those categories.

Most of the study’s patients were female (69%), with a mean age of 58 years. The mean disease duration was not reported, but all patients’ disease durations were less than 1 year. Median baseline DAS was 5.2, but neither mean baseline HAQ nor Sharp scores were reported. All study participants had no history of prior treatment with MTX or csDMARDs in general.

Non-TNF Biologics

Non-TNF Biologic Alone or Plus MTX Versus MTX

Five RCTs compared non-TNF biologics alone or combined with MTX versus MTX monotherapy; each took place over 2 years across multiple countries. Two trials compared combination abatacept (ABA) and MTX versus MTX alone; one of these had a third intervention arm for ABA alone. Another two trials compared combination tocilizumab (TCZ) and MTX versus MTX alone; both had a third intervention arm for TCZ alone. Both were also previously described above in the csDMARDs versus Non-TNF Biologics section. The fifth trial compared different doses of combination rituximab (RIT) and MTX versus MTX alone.

Most of the individuals enrolled in these RCTs were female (67% to 81% across treatment arms), with a mean age between 47 and 54 years. Participants in two trials had average disease durations of approximately 6 months; in another two trials, participants’ average disease durations were about 1 month and 3 months; and participants in the fifth had an average disease duration of approximately 1 year. Across the RCTs, average baseline DAS ranged from 5.2 to 7.1, and average or median baseline HAQ ranged from 1.2 to 1.8. Four of the trials reported average or median baseline Sharp score, which ranged from 5.7 to 7.7 except in one study whose median Sharp score was 0.0. All five trials targeted treatment of aggressive early RA: more than 72 percent of the patients in all five trials were RF seropositive; more than 86 percent in the three trials reported anticyclic citrullinated peptide (anti-CCP) seropositivity were seropositive, and 100 percent in two trials reported erosive disease.

Information about prior treatment for RA was available in four trials. Only one of these trials reported prior csDMARD use, specifically, in about one-third of its patients (30%). All patients enrolled in these four trials were MTX-naïve.

TNF Versus Non-TNF

One RCT (1 year in duration) compared TNF and non-TNF therapies in the United Kingdom. It compared RIT and ADA or ETN and addressed KQs 1, 2, and 3.

The mean age of enrolled individuals was 57 years; a majority were female (72%). The average disease duration in the intervention arms ranged from 6.7 to 8.0 months across treatment arms. The average baseline DAS was 6.2; the median baseline HAQ was 1.7 to 1.8. Baseline Sharp score was not reported. This trial targeted treatment of aggressive early RA: 100 percent of participants were either RF or anti-CPP seropositive.

All study participants had prior MTX use; previous use of csDMARDs in general was not reported at all.
Combinations and Therapy Strategies

We included four RCTs and two observational studies that evaluated combination and therapy strategies. All four trials contributed results to KQs 1, 2, and 3; results in the observational studies were limited to KQ 3 (Appendix C). One trial (five articles) had also been included in the prior report (Table 6).

Four studies enrolled mixed populations of early RA patients and those with longer-duration RA. Only two studies enrolled samples entirely made up of early RA patients with disease duration ≤1 year.

These six studies were conducted in Denmark, France, Ireland and the United Kingdom, the Netherlands, and the United States. The specific combinations and therapy strategies that these researchers compared are described in Appendix C. Study durations ranged from 1 year to 10 years.

Most individuals enrolled in these studies were female (65% to 80%), with a mean age between 46 and 58 years. Two trials reported mean disease duration, which ranged from 2.9 to 4.5 months across treatment arms. The other four studies reported median disease duration, which ranged from 23 weeks to 9 months across treatment arms.

Five studies reported mean or median baseline DAS ranging from 4.3 to 6.2, and they also reported mean or median baseline HAQ ranging from 1.0 to 1.7. Four of these studies reported mean or median baseline Sharp scores ranging from 2.4 to 7.5 across treatment arms. Only one study did not report baseline DAS, HAQ, or Sharp scores. Additionally, a single study targeted treatment of aggressive early RA: 90 percent were RF seropositive and 3 percent were anti-CCP seropositive.

Regarding prior use of MTX, five studies reported at least some information: four of these enrolled only MTX-naïve patients, and only one enrolled some patients with prior MTX treatment (about 20%). As for prior use of csDMARDs in general, all six studies of combination and therapy strategies provided some information. Three studies enrolled samples with any prior csDMARD use, varying greatly from study to study (8.5%, 24%, and 100%), but each of the three remaining studies’ samples was csDMARD naïve.

KQ 1: Comparative Benefits of Drug Therapies for Patients With Early RA in Relation to Disease Activity, Progression of Radiographic Joint Damage, or Remission

Key Points

- Conclusions below are based on early RA studies including patients with moderate to high disease activity, and the majority were MTX naive.
- Higher remission rates were achieved with a combination of corticosteroids plus MTX than with MTX monotherapy (difference in remission ranges from 2.1% to 42.8% over 18 months to 2 years) (low SOE).
- Combination therapy of corticosteroids plus csDMARDs versus csDMARD monotherapy did not differ significantly in disease activity in the long term (up to 5 years) (low SOE).
- Combination therapy of csDMARDs (predominantly MTX plus SSZ) versus csDMARD monotherapy (MTX) did not differ in ACR50 response or remission (low SOE).
- Evidence was insufficient to compare the impact of csDMARD monotherapy versus csDMARD monotherapy.
• The TNF biologic ADA plus MTX had statistically significantly higher ACR50 response (ACR50 difference 22%), smaller radiographic changes (modified Sharp score difference -3.6), and higher remission rates (difference in remission 24%) than ADA monotherapy (moderate SOE).

• The TNF biologics—ADA, CZP, ETN, or IFX—plus MTX had higher remission rates (difference in remission ranges from 5.6% to 70.0% over 26 weeks to 2 years) (low SOE), and two TNF biologics—CZP and ETN—plus MTX had smaller radiographic changes than MTX monotherapy (difference of mTSS change -0.6 to -2.1 over 24 weeks to 2 years) (low SOE for CZP and moderate SOE for ETN). Evidence was insufficient to compare the impact of ADA or IFX plus MTX versus MTX monotherapy for radiographic changes.

• The non-TNF biologics—ABA, RIT, TCZ—plus MTX had smaller radiographic changes (several radiographic measures used) (low SOE for ABA and moderate SOE for RIT, TCZ) and higher remission rates (difference in remission ranges 18% to 38%) (low SOE for TCZ to moderate SOE for ABA, RIT) than MTX monotherapy.

• Evidence was insufficient to determine any differences between one biologic and another biologic for ACR50 response, remission, or radiographic changes.

• With respect to combination therapy, long-term studies show no differences in remission rates or radiographic change between initial combination versus step-up therapies (moderate SOE).

**Detailed Synthesis**

Table 7 presents major findings from trials or other studies used to answer KQ 1 on several intermediate outcomes. It is organized essentially as the syntheses below: corticosteroids; csDMARDs and tsDMARDs; biologics; and drug combinations or other strategies for treating patients with early RA.

Because of the dearth of trials directly comparing interventions of interest, we employed network meta-analyses. For KQ 1, we conducted network meta-analyses on the following outcomes: ACR50 response (13 trials), radiographic joint damage (11 trials), remission (10 trials). For NWMA, we focused on a time period around 1 year (52 to 56 weeks) because data were more comprehensive for this time period than for other ones. For other time points, data were insufficient for NWMA, or clinical heterogeneity across trials was too high to derive meaningful estimates from NWMA. We present results of NWMA on ACR50 and radiographic joint damage within each comparison section below; results on remission are presented in Appendix H. For remission, NWMA rendered mostly inconclusive findings with wide confidence intervals.
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Year, Risk of Bias Rating</th>
<th>Study Design (Duration)</th>
<th>Comparison (Dose)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CAMERA-II, 2012&lt;sup&gt;24&lt;/sup&gt;</td>
<td>RCT (N=239, 2 yrs)</td>
<td>PRED (10 mg/day) + MTX (10 mg/wk) vs. MTX</td>
<td>No significant differences in DAS28, ACR20, ACR50, or remission. Higher ACR70 response at 2 yrs (38.0% vs. 19.0%, mean difference 18.3%, p=0.002) No significant differences in median total SHS scores. Median erosive SHS joint damage less for MTX + PRED vs. MTX (0 [IQR 0 to 0] vs. 0 [IQR 0 to 2], p=0.022)</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CARDERA, 2007&lt;sup&gt;25&lt;/sup&gt;</td>
<td>RCT (N=467, 2 yrs)</td>
<td>PNL (60 mg/day tapered over 34 wks) + MTX (7.5-15 mg/wk) vs. MTX</td>
<td>No significant difference in mean DAS28 change (-1.4 vs. -1.4, p=NR) at 2 yrs DAS28 &lt;2.6 remission (20.0% vs. 17.9%, p=NR) at 2 yrs Lower Larsen score mean change for MTX + PNL vs. MTX (4.7 vs. 7.4, p=0.008) at 2 yrs</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>Todoerti et al., 2010&lt;sup&gt;26&lt;/sup&gt;</td>
<td>RCT, open label (N=210, 2 yrs)</td>
<td>PRED (12.5 mg/day for 1-2 wks then taper to 6.25 mg/day) + MTX (10-20 mg/wk)</td>
<td>Higher DAS &lt;1.6 remission (76.7% vs. 33.3%, p=0.01) at 18 months</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>Montecucco et al., 2012&lt;sup&gt;27&lt;/sup&gt;</td>
<td>RCT, open label (N=220, 1 yr)</td>
<td>PRED (12.5 mg/day for 2 wks then taper to 6.25 mg/day) + MTX (10-25 mg/wk) vs. MTX (10-25 mg/wk)</td>
<td>No significant difference in proportion with low disease activity (80.2% in PRED + MTX vs. 75.5%, p=0.44) at 12 months Higher DAS &lt;2.6 remission (44.8% vs. 27.8%, p=0.02) at 12 months</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CareRA 2015&lt;sup&gt;28&lt;/sup&gt;, 2015&lt;sup&gt;29&lt;/sup&gt;, 2017&lt;sup&gt;30&lt;/sup&gt;</td>
<td>RCT, open label (N=379, 2 yrs)</td>
<td>High-risk patients: MTX (15 mg/wk) + SSZ (2 g/day) + PRED (60 mg/day tapered to 7.5 mg/day) vs. MTX + PRED (30 mg tapered to 5 mg/day) vs. MTX + LEF (10 mg/day) + PRED (30 mg tapered to 5 mg/day) vs. Low-risk patients: MTX (15 mg/wk) vs. MTX + PRED (30 mg tapered to 5 mg/day)</td>
<td>No significant differences in DAS28 change (2.5, 2.3, 2.3, 2.1, 2.1, p=NS) at 52 weeks No significant differences in mean SHS change (0.3, 0.4, 0.3, 0.3, 0.3, p=NS) at 52 weeks</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>BARFOT #2, 2005,118 2014,119 2014,140</td>
<td>RCT, open label N=259 2 yrs (4-yr followup)</td>
<td>PNL 7.5 mg/day + DMARD (SSZ 2 g/day or MTX 10 mg/wk) vs. DMARD (SSZ 2 g/day or MTX 10 mg/wk)</td>
<td>Lower mean DAS28 score in PNL + DMARD vs. DMARD (2.7 vs. 3.2, p=0.005) and higher DAS28 &lt;2.6 remission (55.5% vs. 32.8%, p=0.0005) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td>IDEA, 2014196</td>
<td>RCT N=112 78 wks (1-26 wks blinded, 26-78 wks open label)</td>
<td>IFX (3 mg/kg at wks 0, 2, 6, 14, 22) + MTX (10-20 mg/wk) vs. methyl-PNL (250 mg single dose) + MTX</td>
<td>No differences in ACR50 response (54.0% vs. 55.1%, p=NR) at 26 wks or wk 78 (64.3% vs. 63.4%, p=NR)</td>
</tr>
<tr>
<td></td>
<td>Durez et al., 200719 19a b</td>
<td>RCT N=44 1 yr</td>
<td>IFX (3 mg/kg at wks 0,2,6 until 46 wks) + MTX (7.5-20 mg/wk) vs. methyl-PNL (1 g/wk 0,2,6 and every 8 wks until 46 wks) + MTX vs. MTX</td>
<td>No differences between groups for ACR20, 50, 70 response (p=NR)</td>
</tr>
<tr>
<td></td>
<td>BARFOT #1, 200337</td>
<td>RCT N=245 2 yrs</td>
<td>PNL (7.5-15 mg/day for 1-3 months) + MTX (5-15 mg/wk) vs. SSZ (2-3 g/day) + PNL (up to 10 mg/day)</td>
<td>No significant differences in DAS28 &lt;2.6 remission (29.0% vs. 19.0%, p=0.095) at 2 yrs No significant differences in Larsen score mean change (6.2 vs. 4.1, p=0.298) at 2 yrs</td>
</tr>
<tr>
<td>csDMARD Monotherapy vs. csDMARD Monotherapy</td>
<td>NOR-DMARD, 201228</td>
<td>Observation N=1,102 3 yrs</td>
<td>SSZ (2 g/day) vs. MTX (10-15 mg/wk)</td>
<td>No significant difference in mean DAS28 change for SSZ vs. MTX after adjustment for baseline characteristics (-1.0 vs. -1.5, p=0.71) at 6 months</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Dougados et al., 199911a Maillereft et al., 2003104</td>
<td>RCT N=209 5 yrs</td>
<td>SSZ (2-3 g/day) + MTX (7.5 to 15 mg/wk) vs. SSZ vs. MTX</td>
<td>Significantly decreased change in DAS for SSZ + MTX, compared with SSZ or MTX only (-1.3 vs. -1.1 vs. -0.9, p=0.019) at 1 yr; No significant difference in ACR20 responses (p=NR) No significant changes in DAS at 5 yrs (p=0.9) No significant difference in mTSS change (3.5, 4.6, 4.5, p=NS) at 1 yr or at 5 yrs (p=0.7)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Haagsma et al., 1997&lt;sup&gt;2,2a&lt;/sup&gt;</td>
<td>RCT</td>
<td>SSZ (1-3 g/day) vs. MTX (7.5-15 mg/wk) vs. MTX + SSZ</td>
<td>No significant differences in DAS (-1.6, -1.7, -1.9, p=NS) over 1 yr</td>
</tr>
<tr>
<td></td>
<td>Nijmegen RA Inception Cohort, 2009&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Observational</td>
<td>MTX (7.5-30 mg/wk) vs. SSZ (750-3,000 mg/day) + MTX</td>
<td>No significant differences in DAS28 change after 1 yr between groups (p=0.153)</td>
</tr>
<tr>
<td></td>
<td>COBRA, 1997&lt;sup&gt;24&lt;/sup&gt;, 2002&lt;sup&gt;100&lt;/sup&gt;, 2009&lt;sup&gt;141&lt;/sup&gt;</td>
<td>RCT</td>
<td>PNL (60 mg tapered over 28 wks) + MTX (7.5 mg/wk stopped after 40 wks) + SSZ (2,000 mg/day) vs. SSZ</td>
<td>No significant difference in DAS28 mean change after 5 yrs (-0.02 vs. -0.13, p=0.265)</td>
</tr>
<tr>
<td></td>
<td>COBRA-Light, 2014&lt;sup&gt;25, 110&lt;/sup&gt;</td>
<td>RCT, open label</td>
<td>PNL (60 mg tapered to 7.5 mg/day) + MTX (7.5 mg/wk) + SSZ (1-2g/day) vs. PNL (30 mg tapered to 7.5 mg/day) + MTX (10 mg/d with stepwise increments to 25 mg/week) + ETN intensification in both groups if DAS&gt;1.6 at week 25 or 39</td>
<td>No significant difference in DAS mean changes (1.7, 1.9, p=0.15) over 1 yr</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>FIN-RACO, 1999, 2010, 2013, 2004, 2007, 2010</td>
<td>RCT, open label N=199</td>
<td>MTX (7.5-10 mg/wk) + HCQ (300 mg/day) + SSZ (1 g/day) + PNL (5-7.5 mg/day) vs. DMARD (SSZ 2-3 g/day, which could be changed to MTX 7.5-15 mg/wk if AE or lack of response)</td>
<td>Clinical remission (defined by ACR preliminary criteria) significantly higher in combination group (37.1% vs. 18.4%, p=0.003) at 2 yrs; ACR50 numerically higher in combination group but not significant (71.1% vs. 58.1%, p=0.058) Sustained DAS28 remission at 6 mo, 1 yr, and 2 yrs significantly higher in combination group (OR, 5.6; 95% CI, 2.60-11.55) No significant difference in 5-yr remission (28% vs. 22%, p=NS) Significantly lower Larsen score in combination group (4.0 vs. 12.0, p=0.002) at 2 yrs</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>tREACH, 2013, 2014, 2016</td>
<td>RCT, open label N=515</td>
<td>MTX (25 mg/wk) + SSZ (2 g/day) + HCQ (400 mg/day) + GCs intramuscularly vs. MTX + SSZ + HCQ + GC oral taper (15 mg/day tapers off at 10 wks) vs. MTX + GC oral taper</td>
<td>No significant difference in DAS mean change (-1.8 vs. -1.7 vs. -1.7, p=NR) at 1 yr No significant difference in change in mTSS at 1 yr</td>
</tr>
<tr>
<td>TNF Biologic + csDMARD vs. TNF Biologic</td>
<td>PREMIER, 2006, 2008, 2010, 2010, 2012, 2013, 2014, 2015</td>
<td>RCT</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/wk) vs. ADA vs. MTX</td>
<td>Significantly higher ACR50 in ADA + MTX vs. monotherapies (59.0%, 37.0%, 43.0%, p&lt;0.001) at 2 yrs Significantly higher DAS28 &lt;2.6 remission in ADA + MTX vs. monotherapies (49.0%, 25.0%, 25.0%, p&lt;0.001) at 2 yrs Significantly lower modified Sharp score in ADA + MTX vs. monotherapies (1.9, 5.5, 10.4, p&lt;0.001) at 2 yrs</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic or csDMARD</td>
<td>AVERT, 2015</td>
<td>RCT</td>
<td>ABA (125 mg/wk) + MTX (7.5-20 mg/wk) vs. ABA vs. MTX</td>
<td>DAS28 &lt;2.6 remission significantly highest in ABA + MTX (60.9%, 42.5%, 45.2%, p=0.010 for ABA + MTX vs. MTX) at 1 yr</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design, N, Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic or csDMARD</td>
<td>FUNCTION, 2016&lt;sup&gt;12&lt;/sup&gt; a 2017&lt;sup&gt;13-14&lt;/sup&gt; a d</td>
<td>RCT, N=1,162 2 yr Aggressive RA</td>
<td>TCZ (4 mg/kg monthly) + MTX (7.5-20 mg/wk) vs. TCZ (8 mg/kg monthly) + MTX vs. TCZ (8 mg/kg) vs. MTX</td>
<td>Significantly higher ACR50 response rates for TCZ + MTX vs. MTX (54.9%, 56.2%, 50.7%, 41.5%, p&lt;0.014) at 1 yr; similar findings (36.5%, 57.6%, 53.1%, 22.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher DAS28-ESR remission for TCZ 8 mg + MTX vs. MTX (34.0%, 49.0%, 39.4%, 19.5%, p&lt;0.0001) at 1 yr; similar findings (28.1%, 47.6%, 43.5%, 16.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lowest radiographic mTSS score change for TCZ 8 mg + MTX (0.4, 0.1, 0.3, 1.1, p=0.0001) at 1 yr; similar findings (1.4, 0.2, 0.6, 1.9, p=NR) at 2 yrs</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic or csDMARD</td>
<td>U-Act-Early, 2016&lt;sup&gt;11&lt;/sup&gt; a d</td>
<td>RCT, N=317 2 yrs</td>
<td>TCZ (8 mg/kg monthly) + MTX (10-30 mg/wk) vs. TCZ vs. MTX</td>
<td>No significant differences in median DAS change (3.3, 3.3, 3.2, p=0.66) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Higher DAS28 remission with TCZ + MTX and TCZ arms than MTX (86.0% vs 83.0% vs 48.0%, p &lt;0.001 ) at 24 weeks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Higher DAS remissions with TCZ + MTX and TCZ arms than MTX (86.0% vs. 88.0% vs. 77.0%, p=0.036 for TCZ vs. MTX, p=0.06 for TCZ + MTX vs. MTX) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly lower radiographic SHS mean change from baseline with TCZ + MTX (1.2, 1.4, 1.5, p=0.06 for TCZ vs. MTX, p=0.016 for TCZ + MTX vs. MTX) at 2 yrs</td>
</tr>
<tr>
<td>csDMARDs vs. tsDMARDs</td>
<td>Conaghan et al., 2016&lt;sup&gt;29&lt;/sup&gt;</td>
<td>RCT, N=108 1 yr</td>
<td>TOF (20 mg/day) + MTX (10-20 mg/wk) vs. TOF vs. MTX</td>
<td>Significantly higher DAS28-4 ESR &lt;3.2 in TOF + MTX vs. monotherapies (58.8%, 30.6%, 18.9%, p&lt;0.001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher ACR50 response in TOF + MTX (65.7%, 50.0%, 35.1%, p&lt;0.01) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher DAS28-4 ESR &lt;2.6 remission in TOF + MTX (35.3%, 19.4%, 13.5%, p&lt;0.05) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly smaller change in radiographic mTSS for TOF (-0.1) compared with TOF + MTX (0.8) and MTX (1.4) (p&lt;0.05) at 1 yr</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design (N Duration)</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>HIT HARD, 2013, Medium (DAS, ACR)</td>
<td>RCT, open label (N=172; 48 wks)</td>
<td>ADA (40 mg biwkly x 24 wks) + MTX (15 mg/wk) vs. MTX</td>
<td>No significant differences in DAS (3.2 vs. 3.4, p=0.41) or ACR50 response (52.6% vs. 51.4%, p=0.88) at 48 wks. No significant differences in DAS remission (42.4% vs. 36.8%, p=0.47) at 48 wks. Significantly less radiographic mTSS change for ADA + MTX (2.6 vs. 6.4, p=0.01) at 48 wks.</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>HOPEFUL 1, 2014, Medium</td>
<td>RCT (N=334; 26 wks (plus 6-month open label))</td>
<td>ADA (40 mg biwkly) + MTX (6-8 mg/wk) vs. MTX</td>
<td>Numerically higher ACR50 with ADA + MTX vs. MTX (64.3% vs. 38.7%, p=NR) at 26 wks. Significantly higher DAS &lt;2.6 remission with ADA + MTX vs. MTX (31.0% vs. 14.7%, p&lt;0.001) after 26 wks. Significantly less radiographic mTSS mean change with ADA + MTX vs. MTX (1.5 vs. 2.4, p&lt;0.001) at 26 wks.</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>OPTIMA, 2013, 2014, 2016, Low</td>
<td>RCT (N=1,032; 78 wks (open label after 26 wks))</td>
<td>ADA (40 mg biwkly) + MTX (7.5-20 mg/wk) vs. ADA vs. MTX</td>
<td>Significantly higher ACR50 for ADA + MTX vs. monotherapies (59.0%, 37.0%, 43.0%, p&lt;0.001) at 2 yrs. Significantly higher DAS &lt;2.6 remission in ADA + MTX vs. MTX (34.0% vs. 17.0%, p&lt;0.001) at 26 wks. Significantly lower modified Sharp score in ADA + MTX vs. monotherapies (1.9, 5.5, 10.4, p&lt;0.001) at 2 yrs.</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>PREMIER, 2006, 2008, 2010, 2012, 2013, 2014, 2015, Medium</td>
<td>RCT (N=799; 2 yrs)</td>
<td>ADA (40 mg biwkly) + MTX (7.5-20 mg/wk) vs. ADA vs. MTX</td>
<td>Significantly higher ACR50 in ADA + MTX vs. monotherapies (59.0%, 37.0%, 43.0%, p&lt;0.001) at 2 yrs. Significantly higher DAS &lt;2.6 remission in ADA + MTX vs. monotherapies (49.0%, 25.0%, 25.0%, p&lt;0.001) at 2 yrs. Significantly lower modified Sharp score in ADA + MTX vs. monotherapies (1.9, 5.5, 10.4, p&lt;0.001) at 2 yrs.</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>PROWD, 2008, 2016, Medium (16-wk outcomes)</td>
<td>RCT (N=148; 56 wks)</td>
<td>ADA (40 mg biwkly) + MTX (7.5-25 mg/wk) vs. MTX</td>
<td>No significant differences in ACR50 (56.0% vs. 45.2%, p=0.189) at 56 wks. No significant differences in DAS28 &lt;2.6 remission (48.0% vs. 36.1%, p=0.145) at 56 wks.</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>------------------------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>C-OPERA, 2016&lt;sup&gt;13&lt;/sup&gt;, 2017&lt;sup&gt;153a&lt;/sup&gt;</td>
<td>RCT N=316 2 yrs Aggressive RA</td>
<td>CZP (400 mg biwkly x 4 wks, then 200 mg biwkly) + MTX (8-12 mg/wk) vs. MTX</td>
<td>Significantly higher DAS28-ESR remission for CZP + MTX vs. MTX (52.8% vs. 30.6%, p&lt;0.001) at 24 wks; no significant differences (41.5% vs. 33.1%, p=0.132) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td>Medium (24 wks) High (52 wks, 2 yrs)</td>
<td></td>
<td></td>
<td>Significantly lower radiographic mTSS mean change for CZP + MTX vs. MTX (0.3 vs. 0.9, p=0.003) at 24 wks; similar findings (0.7 vs. 3.0, p=0.001) at 2 yrs</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>C-EARLY 2017&lt;sup&gt;38, 39a&lt;/sup&gt;</td>
<td>RCT N=879 52 wks Aggressive RA</td>
<td>CZP (400 mg biwkly) + MTX (10-25 mg/wk) vs. MTX</td>
<td>Significantly higher ACR50 for CZP + MTX vs. MTX (61.8% vs. 52.6%, p=0.023) at 52 wks</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td>Significantly higher DAS28-ESR remission for CZP +MTX vs. MTX (42.6% vs. 26.8%, p&lt;0.001) at 52 wks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant radiographic mTSS change from baseline for CZP + MTX vs. MTX (70.3% vs. 49.7%, p&lt;0.001) at 52 wks</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>COMET, 2008&lt;sup&gt;12&lt;/sup&gt;, 2009&lt;sup&gt;154&lt;/sup&gt;, 2010&lt;sup&gt;108, 109&lt;/sup&gt;, 2012&lt;sup&gt;155&lt;/sup&gt;, 2014&lt;sup&gt;156 a&lt;/sup&gt;</td>
<td>RCT N=542 2 yrs</td>
<td>ETN (50 mg/wk) + MTX (7.5-20 mg/wk) vs. MTX</td>
<td>Significantly higher ACR50 response for ETN + MTX vs. MTX (70.7% vs. 49.0%, p&lt;0.0001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td>Significantly improved DAS &lt;1.6 remission for ETN + MTX vs. MTX (51.3% vs. 27.8%, p&lt;0.0001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Numerically lower radiographic mTSS change for ETN + MTX vs. MTX (0.3, 2.4, p=NR) at 1 yr</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Enbrel ERA, 2000&lt;sup&gt;14&lt;/sup&gt;, 2002&lt;sup&gt;110&lt;/sup&gt;, 2005&lt;sup&gt;112&lt;/sup&gt;, 2006&lt;sup&gt;111 a&lt;/sup&gt;</td>
<td>RCT N=632 1 yr (1-yr open label extension) Aggressive RA</td>
<td>ETN (25 mg twice wkly) vs. MTX (7.5-20 mg/wk)</td>
<td>No significant difference in ACR20 response rates (65.0% vs. 72.0%, p=0.16) at yr 1</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td>Significantly higher ACR20 response for ETN than MTX (72.0% vs. 59.0%, p=0.005) at yr 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant difference in radiographic mean mTSS change (1.6 vs. 1.0, p=0.11) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly lower radiographic mTSS mean change for ETN than MTX (1.3 vs. 3.2, p=0.001) at 2 yrs</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Marcora et. al, 2006&lt;sup&gt;113&lt;/sup&gt;</td>
<td>RCT N=26 26 wks</td>
<td>ETN (25 mg twice wkly) vs. MTX (7.5-15 mg/wk)</td>
<td>No significant difference in DAS28 (3.2 vs. 3.1, p=0.53) at 24 wks</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design (N)</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>ASPIRE, 2004,17 2006,107 2009,106 2017157 a</td>
<td>RCT N=1,049 54 wks</td>
<td>IFX (3 mg/kg/8 wks) + MTX (20 mg/wk) vs. IFX (6 mg/kg/8 wks) + MTX vs. MTX</td>
<td>Significantly higher ACR50 response in both IFX + MTX groups vs. MTX (45.6% vs. 50.4% vs. 32.1%, p&lt;0.001) at 54 wks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher remission (DAS28-ESR &lt;2.6) for IFX + MTX vs. MTX groups combined (21.3% vs. 12.3%, p&lt;0.001) at 54 wks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly lower radiographic mTSS score changes in both IFX + MTX groups vs. MTX (0.4, 0.5, 3.7, p&lt;0.001) at 54 wks</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>Quinn et al., 200541 a</td>
<td>RCT N=20 2 yrs</td>
<td>IFX (3 mg/kg at 0, 2, 6, and every 8 wks) + MTX (7.5-25 mg/wk) vs. MTX</td>
<td>Numerically higher ACR50 response but not significant (70.0% vs. 50.0%, p=NS) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Higher remission for IFX + MTX vs. MTX (70.0% vs. 20.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant change in radiographic mean SHS scores (10.0 vs. 12.0, p=NR) at 2 yrs</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>Durez et al., 200718 a b</td>
<td>RCT N=44 1 yr</td>
<td>IFX (3 mg/kg at wks 0,2,6 until 46 wks) + MTX (7.5-20 mg/wk) vs. MTX</td>
<td>No differences between groups for ACR20, 50, and 70 response (p=NR) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No differences between groups for DAS28-CRP (2.8 vs. 3.3, p=NR) at 1 yr</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Combination Therapy</strong></td>
<td>IMPROVED, 2013,9 2014,158 2016120</td>
<td>RCT N=161 2 yrs</td>
<td>ADA (40 mg biwkly) + MTX (25 mg/wk) vs. MTX + PRED (7.5 mg/day) + HCQ (400 mg/day) + SSZ (2 g/day)</td>
<td>No significant differences in DAS or DAS &lt;1.6 remission at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in radiographic mTSS score progression (6.4% vs. 10.8%, p=0.31) at 2 yrs</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Combination Therapy</strong></td>
<td>SWEFOT, 2009,10 2012,122 2013,121, 123, 126 2015125 2016124</td>
<td>RCT, open label N=258 1 yr (2-yr followup)</td>
<td>IFX (3 mg/kg at 0,2,6 weeks then every 8 wks) + MTX (20 mg/wk) vs. MTX + SSZ (2 g/day) + HCQ (400 mg/day)</td>
<td>Significantly higher ACR50 response for IFX + MTX vs. MTX + SSZ + HCQ (25.0% vs. 14.6%, p=0.0424) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in ACR50 or ACR70 responses or remission at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in SHS scores at 5-yr followup</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>AGREE, 2009, 2011, 2015</td>
<td>RCT N=509 1 yr (1-yr open-label extension)</td>
<td>ABA (10 mg/kg on days 1, 15, and 29 and every 4 wks after) + MTX (7.5-20 mg/wk) vs. MTX</td>
<td>Significantly reduced DAS28 activity for ABA + MTX vs. MTX (-3.2 vs. -2.5, p&lt;0.001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher ACR50 response rates for ABA + MTX vs. MTX (57.4 vs. 42.3%, p&lt;0.001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher remission rates for ABA + MTX than MTX (41.4% vs. 23.3%, p&lt;0.001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly less mean radiographic changes by Genant-modified Sharp score (0.6 vs. 1.1, p=0.040) at 1 yr</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>AVERT, 2015</td>
<td>RCT N=351 2 yrs</td>
<td>ABA (125 mg/wk) + MTX (7.5-20 mg/wk) vs. ABA vs. MTX</td>
<td>DAS28 &lt;2.6 remission significantly highest in ABA + MTX (60.9%, 42.5%, 45.2%, p=0.010 for ABA + MTX vs. MTX) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher ACR50 response rates for TCZ + MTX vs. MTX (54.9%, 56.2%, 50.7%, 41.5%, p&lt;0.014 at 1 yr; similar findings (36.5%, 57.6%, 53.1%, 22.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher DAS28-ESR remission for TCZ 8 mg + MTX vs. MTX (34.0%, 49.0%, 39.4%, 19.5%, p&lt;0.0001) at 1 yr; similar findings (28.1%, 47.6%, 43.5%, 16.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lowest radiographic mTSS score change for TCZ 8 mg vs. MTX (0.4, 0.1, 0.3, 1.1, p=0.0001) at 1 yr; similar findings (1.4, 0.2, 0.6, 1.9, p=NR) at 2 yrs</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>FUNCTION, 2016</td>
<td>RCT N=1,162 2 yr</td>
<td>TCZ (4 mg/kg monthly) + MTX (7.5-20 mg/wk) vs. TCZ (8 mg/kg monthly) + MTX vs. TCZ (8 mg/kg) vs. MTX</td>
<td>Significantly higher ACR50 response rates for TCZ + MTX vs. MTX (54.9%, 56.2%, 50.7%, 41.5%, p&lt;0.014) at 1 yr; similar findings (36.5%, 57.6%, 53.1%, 22.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher DAS28-ESR remission for TCZ 8 mg + MTX vs. MTX (34.0%, 49.0%, 39.4%, 19.5%, p&lt;0.0001) at 1 yr; similar findings (28.1%, 47.6%, 43.5%, 16.0%, p=NR) at 2 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lowest radiographic mTSS score change for TCZ 8 mg vs. MTX (0.4, 0.1, 0.3, 1.1, p=0.0001) at 1 yr; similar findings (1.4, 0.2, 0.6, 1.9, p=NR) at 2 yrs</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>IMAGE, 2011</td>
<td>RCT N=755 1 yr</td>
<td>RIT (1 g days 1 and 15) + MTX (7.5-20mg/wk) vs. MTX (7.5-30 mg/wk) vs. RIT (500 mg days 1 and 15) + MTX vs. MTX</td>
<td>Significantly higher rate of low disease activity (DAS28) in RIT + MTX groups vs. MTX (43.0%, 40.0%, 20.0%, p&lt;0.001) at 1 yr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly higher remission (DAS &lt;2.6) in RIT + MTX groups vs. MTX (31.0%, 25.0%, 13.0%, p&lt;0.0010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significantly less radiographic change in RIT + MTX groups vs. MTX by Genant-modified Sharp (0.4, 0.6, 1.1, p&lt;0.0001)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>U-Act-Early, 2016&lt;sup&gt;11&lt;/sup&gt; Medium</td>
<td>RCT N=317 2 yrs</td>
<td>TCZ (8 mg/kg monthly) + MTX (10-30 mg/wk) vs. TCZ vs. MTX</td>
<td>No significant differences in median DAS change (3.3, 3.3, 3.2, p=0.66) at 2 yrs. Higher DAS28 remission with TCZ + MTX and TCZ arms than MTX (86.0% vs 83.0% vs 48.0%, p &lt;0.001 ) at 24 weeks. Higher DAS remissions with TCZ + MTX and TCZ arms than MTX (86.0% vs. 88.0% vs. 77.0%, p=0.036 for TCZ vs. MTX, p=0.06 for TCZ + MTX vs. MTX) at 2 yrs. Significantly lower radiographic SHS mean change from baseline with TCZ + MTX (1.2, 1.4, 1.5, p=0.06 for TCZ vs. MTX, p=0.016 for TCZ + MTX vs. MTX) at 2 yrs.</td>
</tr>
<tr>
<td>TNF vs. Non-TNF</td>
<td>ORBIT, 2016&lt;sup&gt;8&lt;/sup&gt; High</td>
<td>RCT N=329 1 yr</td>
<td>RIT (1 g days 1 and 15 and after 26 wks if persistent disease activity) vs. ADA (40 mg biwkly) or ETN 50 mg/wk</td>
<td>No significant differences in DAS28-ESR (-2.6 vs.-2.4, p=0.24) at 1 yr.</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>BeSt, 2005,&lt;sup&gt;9&lt;/sup&gt;, 2007,&lt;sup&gt;5&lt;/sup&gt;, 2008,&lt;sup&gt;4&lt;/sup&gt;, 2009,&lt;sup&gt;3, 8&lt;/sup&gt;, 2010,&lt;sup&gt;1&lt;/sup&gt;, 2011,&lt;sup&gt;99, 90&lt;/sup&gt;, 2012,&lt;sup&gt;90, 91&lt;/sup&gt;, 2013,&lt;sup&gt;82&lt;/sup&gt;, 2014,&lt;sup&gt;88&lt;/sup&gt;, 2016&lt;sup&gt;37&lt;/sup&gt; Low Medium (10 yr outcomes)</td>
<td>RCT N=508 12 months (10 yr follow-up)</td>
<td>DAS-driven treatment; 1: sequential monotherapy starting with MTX (15 mg/wk) vs. 2: stepped up-combination therapy (MTX, then SSZ, then HCQ, then PRED) vs. 3: combination with tapered high-dose PRED (60 mg/d to 7.5 mg/d) vs. 4: combination (MTX 25-30 mg/wk) with IFX (3 mg/kg every 8 wks, per DAS, could be titrated to 10 mg/kg)</td>
<td>After 1 yr, DAS &lt;2.4: 53.0%, 64.0%, 71.0%, 74.0%; p=0.004 for 1 vs. 3; p=0.001 for 1 vs. 4; p=NS for other comparisons. Shorter time to DAS &lt;2.4 for initial combination therapy groups (groups 3 and 4) than monotherapy groups (groups 1 and 2) (median months; 3, 3, 9, 9; p&lt;0.001) at 2 yrs. No significant differences in remission among groups (DAS &lt;1.6; 50.0%, 41.0%, 38.0%, 42.0%; p=0.40) at 4 yrs. No significant differences in drug-free remission (14.0%, 16.0%, 10.0%, 19.0; p=0.18) at 5 yrs. No significant differences in DAS &lt;1.6 remission (51.0%, 49.0%, 53.0%, 53.0%; p=0.94) at 10 yrs.</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Combination and Therapy Strategies (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>TEAR, 2012,20</td>
<td>RCT N=755 2 yr</td>
<td>Immediate MTX (20 mg/wk) plus ETN (50 mg/wk) vs. Immediate MTX plus SSZ (1-2 g/day) plus HCQ (400 mg/day) vs. Step up MTX to combo (MTX plus ETN) vs. Step up MTX to combo (MTX plus SSZ plus HCQ)</td>
<td>After 4 yrs, significantly less radiographic joint damage in groups 3 and 4 (median SHS change: 5.0, 5.5, 3.0, 2.5; p&lt;0.01 for 1 and 2 vs. 4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>After 5 yrs, significantly less radiographic joint damage in groups 3 and 4 (median SHS change: 2.5, 2.3, 1.0, 1.0; p&lt;0.01 for 1 and 2 vs. 4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>After 10 yrs, no significant differences in radiographic joint damage (mTSS: 11.0, 8.0, 8.0, 6.0; p=0.15)</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>GUEPARD, 200922</td>
<td>RCT N=65 1 yr</td>
<td>1: ADA 40 mg every 2 wks plus MTX; treatment adjusted every 3 mos to achieve DAS28 &lt;3.2 2: MTX (max 20 mg/wk)</td>
<td>ACR50 response higher in ADA + MTX group at 12 wks (84.0% vs. 60.0%, p=NR), but no significant difference at 52 wks (67.0% vs. 68.0%, p=NS, NR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in DAS remission (39.4% vs. 59.4%, p=0.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in radiographic changes (mTSS 1.9 vs. 1.8, p=0.18)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Year, Risk of Bias Rating</td>
<td>Study Design</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>OPERA, 2013, 2014, 2015, 2016, 2017</td>
<td>RCT, open label after yr 1</td>
<td>N=180 2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/wk) vs. MTX</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Included in network meta-analysis.  
*b This study evaluates comparisons in both the High-Dose Corticosteroid and TNF Biologic categories.  
*c This study evaluates comparisons in both the csDMARD and TNF Biologic categories.  
*d These studies evaluate comparisons in both the csDMARD and Non-TNF Biologic categories.

ABA = abatacept; ACR20/50/70 = American College of Rheumatology 20%, 50% and 70% improvement; ADA = adalimumab; AE = adverse event; biwkly = biweekly; csDMARD = conventional synthetic DMARD; CZP = certolizumab pegol; DAS = Disease Activity Score (based on 44 joints); DAS28-ESR = Disease Activity Score 28 using erythrocyte sedimentation rate; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; FIN-RACo = Finnish Rheumatoid Arthritis Combination Therapy trial; g = grams; GC = glucocorticoid; HCQ = hydroxychloroquine; IFX = infliximab; kg = kilogram; LEF = leflunomide; Methyl-PNL = methylprednisolone; mg = milligrams; mTSS = modified Total Sharp/van der Heijde score; MTX = methotrexate; N = number; NR = not reported; NS = not significant; PNL = prednisolone; PRED = prednisone; RA = rheumatoid arthritis; RCT = randomized controlled trial; RIT = rituximab; SD = standard deviation; SHS = Sharp/van der Heijde Score; SSZ = sulfasalazine; TCZ = tocilizumab; TNF = tumor necrosis factor; TOF = tofacitinib; vs. = versus; wk = week; yr = year.

Figure 3 and Figure 4 depict the network diagrams for ACR50 and radiographic joint damage, and Table 8 lists the studies we used in our NWMA of both outcomes. The network structure for both outcomes is mostly “star-shaped” indicating a dearth of head-to-head studies directly comparing interventions. Most effect estimates, therefore, were derived from indirect comparisons relative to MTX, rather than mixed treatment comparisons.
Figure 3. Network diagram for network meta-analysis: ACR50 response rates

ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; N = number of patients.

Figure 4. Network diagram for network meta-analysis: change from baseline in radiographic joint damage score

MTX = methotrexate; N = number of patients.
### Table 8. Studies included in KQ 1 network meta-analyses

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>Study Name</th>
<th>ACR50</th>
<th>Radiographic Joint Damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA + MTX vs. ABA vs. MTX</td>
<td>AVERT, 2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADA + MTX vs. MTX</td>
<td>PROWD, 2008, 2016</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>C-EARLY, 2017</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>C-OPERA, 2016</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IFX + MTX vs. methyl-PNL + MTX vs. MTX</td>
<td>Durez et al., 2007</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>IFX + MTX vs. MTX</td>
<td>Quinn et al., 2005</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>SSZ + MTX vs. SSZ vs. MTX</td>
<td>Dougados et al., 1999, Maillefert et al., 2003</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ vs. MTX</td>
<td>FUNCTION, 2016, 2017</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ vs. MTX</td>
<td>U-Act-Early, 2016</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

a All data used in NWMA were measured at the 1-year followup time point.
b NWMA of DAS remission are presented in Appendix H.
c Outcomes from these studies at the 1-year followup time point were rated as high ROB, and we therefore only used their data in sensitivity analyses presented in Appendix I.

ABA = abatacept; ACR50 = American College of Rheumatology 50% improvement; ADA = adalimumab; CZP = certolizumab pegol; DAS = Disease Activity Score; ETN = etanercept; IFX = infliximab; KQ = Key Question; methyl-PNL = methylprednisolone; MTX = methotrexate; NA = not applicable; NWMA = network meta-analysis; PROWD = PRevention of Work Disability trial; ROB = risk of bias; SSZ = sulfasalazine; TCZ = tocilizumab; vs. = versus.

### Corticosteroids

#### Corticosteroids Versus csDMARDs

Six trials compared the combination of a corticosteroid plus a csDMARD with a csDMARD monotherapy (N=210 to 467) ( }
Table 8). Study durations ranged from 1 to 2 years of active treatment; four were open label trials and all were medium ROB, except one whose 4-year followup data had a high ROB. Treatment arms differed significantly at baseline in terms of patients’ age in one trial, but its statistical analyses adjusted for age as a covariate. In another two trials, baseline similarity between arms was unclear. The csDMARD under examination was MTX in five trials; one study included SSZ; studies did not report any prior history of MTX use. Overall, improvements in disease activity and ACR responses were mixed regarding statistical significance, but they trended toward favoring the treatment combination of corticosteroid plus csDMARD over csDMARD monotherapy. The combination of a corticosteroid plus a csDMARD (SSZ or MTX) demonstrated less radiographic progression in most studies measuring this outcome compared with csDMARD monotherapy. These positive findings were apparent in studies with longer duration (2 years). Additionally, trials ranging from 1 to 2 years of active treatment had significantly higher remission rates with the combination of a corticosteroid plus MTX than MTX monotherapy (remission rates ranging from 44.8% to 76.7% for combination therapy and 27.8% to 33.3% for MTX monotherapy). Overall, higher remission rates were achieved with a combination of corticosteroids plus MTX than MTX monotherapy (low SOE).

High-Dose Corticosteroids

Two trials evaluated the efficacy of high-dose corticosteroids in MTX-naïve populations. Both were medium ROB, and in one trial, baseline characteristics were similar between treatment arms, and although characteristics differed significantly between arms in the other, sensitivity analyses confirmed that those differences had no effect on its findings. The IDEA trial compared the combination of IFX plus MTX with high-dose methylprednisolone (methyl-PNL) plus MTX (N=112). In it, a single high dose of methyl-PNL (250 mg) plus MTX was compared with IFX plus MTX over 26 weeks with a 50-week open-label extension. No significant differences were found in ACR50 responses (disease activity) at 26 or 78 weeks, although response rates were high in both groups (64.3% vs. 63.4% at 78 weeks, p=NR). The two groups did not differ statistically in radiographic changes.

Similarly, a study comparing IFX plus MTX versus high-dose methyl-PNL plus MTX versus MTX monotherapy (N=44) found no significant differences between groups in DAS28-CRP, ACR20, ACR50, and ACR70 responses. In this study, methyl-PNL was dosed at 1g IV at weeks 0, 2, and 6 and then every 8 weeks for 46 weeks. DAS remission was achieved in 40 percent of MTX-treated patients and 70 percent of the methyl-PNL plus MTX group and IFX plus MTX group but without significant differences (p=NR). Radiographic changes were only measured by MRI-detected erosions. There was more significant progression in MRI-detected erosions in the methyl-PNL group compared with patients treated with IFX plus MTX (p=0.035). Overall, the SOE was insufficient for comparisons of high-dose corticosteroid plus MTX therapy with IFX plus MTX.
csDMARDs

csDMARDs Versus csDMARDs

**csDMARD Monotherapy Versus csDMARD Monotherapy**

One 2-year trial (N=245) examined SSZ plus prednisolone versus MTX plus prednisolone and found no statistically significant differences in remission rates (defined by a DAS28 < 2.6) or Larsen score change from baseline (6.2 vs. 4.1, p=0.29). Similarly, one 3-year observational study (n=1,102) compared SSZ with MTX and found no statistically significant differences in mean DAS28 after adjusting for baseline characteristics (-1.04 vs. -1.52, p=0.71). Both studies in MTX-naïve populations were rated high ROB because of high attrition rates, and in one trial, statistically significant baseline differences between treatment arms in RF-positivity and radiographic damage were not accounted for in statistical analyses. Overall, the SOE was insufficient for comparisons between csDMARD monotherapies.

**csDMARD Combination Therapy Versus csDMARD Monotherapy**

Combination therapy with csDMARDs versus csDMARD monotherapy did not differ significantly in disease activity in the long term (up to 5 years) (low SOE). Six trials compared SSZ plus MTX with csDMARD monotherapy (MTX or SSZ) (overall N=1347). Study duration ranged from 1 to 5 years and did not report any prior history of MTX use. Randomization within each of these trials was successful in ensuring the similarity of baseline characteristics between treatment arms, although baseline similarity in one trial was unclear with regard to DAS and Sharp scores. All trials found no significant differences in disease activity at 1 to 5 years. Radiographic changes were consistent but imprecise: two trials reported decreased radiographic progression in the combination therapy arms (two csDMARDs [SSZ plus MTX] or three csDMARDs [SSZ plus MTX plus HCQ plus prednisolone]) compared with monotherapy, another two trials did not find any radiologic differences but trended in favor of combination therapy, and one trial found no radiologic differences between combination therapy and monotherapy without a trend in favor of either.

The observational study (n=230) examined the effect of switching to or adding MTX after patients have attempted SSZ. These patients were switched to MTX (7.5 mg-30 mg/week) or continued on SSZ and MTX was added. After 1 year, these groups did not differ significantly in disease activity.

**csDMARDs Versus Biologics**

**TNF Biologic: MTX Plus TNF Biologic Versus Monotherapy With Either MTX or TNF Biologic**

One RCT provided evidence for direct comparison of a TNF biologic plus MTX versus MTX or TNF biologic monotherapies. The PREMIER study (N=799) compared MTX (20 mg/week) plus the TNF biologic ADA (40 mg biweekly) with either drug alone in MTX-naïve patients with early aggressive RA (8 or more swollen joints, 10 or more tender joints, elevated sedimentation rate or C-reactive protein, rheumatoid factor positive, or at least one joint erosion). ADA plus MTX had significantly higher ACR50 response, smaller radiographic changes, and higher remission rates than ADA monotherapy (moderate SOE). Significantly more patients on MTX plus ADA achieved an ACR50 response than did patients receiving monotherapy with either MTX or ADA (59%, 43%, 37%, p<0.001) at 2 years. Patients in the ADA plus MTX
group had also higher remission rates (49%, 25%, 25%, p<0.001). Additionally, the combination therapy group had lower radiographic progression (modified Sharp/van der Heijde score [mTSS]: 1.9, 5.5, 10.4; p<0.001). During the 10-year open-label extension, patients taking ADA plus MTX had significantly less radiographic progression than those on monotherapy, but results were limited by a 34 percent overall attrition rate.

Results of the NWMA were consistent with the findings of the PREMIER study and favored the combination of MTX plus ADA versus ADA monotherapy for higher ACR50 response (relative risk [RR], 1.52; 95% confidence interval [CI], 1.28 to 1.80) and less radiographic progression (standardized mean difference [SMD], -0.38; 95% CI, -0.55 to -0.21) (Figure 5 for ACR50 and Figure 6 for radiographic joint damage). NWMA also favored the combination of MTX plus ETN versus ETN for higher ACR50 response (RR, 1.57; 95% CI, 1.23 to 2.02) (Figure 5). No comparisons were available for CZP, golimumab (GOL), or IFX. For ACR50 data and radiographic joint damage, Figure 5 and Figure 6 show the forest plots. The network structure for both outcomes is mostly “star-shaped,” indicating a dearth of head-to-head studies directly comparing interventions. Most effect estimates, therefore, were derived from indirect comparisons relative to MTX rather than mixed treatment comparisons.

**Figure 5. Forest plot for network meta-analysis of ACR50 response rates: MTX plus TNF biologic versus TNF biologic**

![Forest plot](https://example.com/forest_plot.png)

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
Non-TNF Biologic: MTX Plus Non-TNF Biologic Versus Monotherapy With Either MTX or Non-TNF Biologic

One RCT, the multinational AVERT study (n=351), compared the combination of MTX (7.5 mg/week) plus ABA (125 mg/week) with ABA monotherapy and also MTX monotherapy (prior MTX use not reported). This double-blind RCT compared treatments over 1 year; at year 2, patients with DAS28-CRP <3.2 were tapered off treatment. If patients experienced an RA flare by month 15, they were given MTX plus ABA. At 1-year (before treatment was withdrawn), patients in the MTX plus ABA group had significantly higher remission (DAS<2.6: 60.9% vs. 42.5% vs. 45.2%, p=0.010) rates than the MTX-only comparison group. Remission rates remained higher for MTX plus ABA than for MTX monotherapy groups following withdrawal at 18 months (14.8% vs. 7.8%, p=0.045).

Two RCTs assessed differences in efficacy between an MTX plus TCZ combination and either MTX or TCZ monotherapy in MTX-naïve populations. MTX plus the non-TNF biologic TCZ led to smaller radiographic changes (low SOE) and higher remission rates than MTX monotherapy (moderate SOE). The FUNCTION trial examined an MTX plus TCZ combination over 1 year in 1,162 patients with early aggressive RA (moderate to severe active RA classified by ACR criteria). After 1 year, 49 percent in the MTX plus TCZ (8 mg/kg/month) combination, 19.5 percent in the MTX monotherapy, and 39.4 percent in the TCZ monotherapy group achieved remission (p<0.001) (low SOE). Similar findings were noted for the FUNCTION trial at 2 years, but this trial was rated high ROB because of high overall attrition. The U-Act-Early trial examined 317 patients with early RA over 2 years. Patients were randomized to MTX (10-30 mg/week) plus TCZ (8 mg/kg/month), MTX monotherapy, and TCZ monotherapy. At the primary outcome time point of 24 weeks, MTX plus TCZ and TCZ monotherapy led to higher DAS28 remission than MTX (86% vs. 83% vs. 48%, p<0.001). MTX plus TCZ and TCZ monotherapy also trended toward higher remission at 2 years than MTX, but the difference was not significant (86% vs. 88% vs. 77%, respectively, p=0.06). Both trials reported less radiographic progression with MTX plus TCZ than with MTX monotherapy.

NWMA favored the combination of MTX plus TCZ over TCZ monotherapy for ACR50 response but was not statistically significant (RR, 1.08; 95% CI, 0.96 to 1.21) (Figure 7), and there were no significant differences in radiographic progression (SMD, -0.03; 95% CI, -0.17 to 0.11).
Similarly, the combination of MTX plus ABA was favored over ABA for ACR50 response, but the difference was not statistically significant (RR, 1.18; 95% CI, 0.95 to 1.47) (Figure 7). No comparisons were available for RIT or sarilumab (SAR).

**Figure 7. Forest plot for network meta-analysis of ACR50 response rates: MTX plus non-TNF versus non-TNF biologic**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. Abatacept</td>
<td>1.18 (0.95, 1.47)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Tocilizumab</td>
<td>1.08 (0.96, 1.21)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

**Figure 8. Forest plot for network meta-analysis of change from baseline in radiographic joint damage score: MTX plus non-TNF versus non-TNF biologic**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tocilizumab+MTX vs. Tocilizumab</td>
<td>-0.03 (-0.17, 0.11)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; SMD = standardized mean difference (mean difference divided by standard deviation); TNF = tumor necrosis factor; vs. = versus.

**csDMARDs Versus tsDMARDs: MTX Plus tsDMARD Versus Either MTX or tsDMARD**

One RCT (n=109) compared the combination of tofacitinib (TOF, 10 mg twice daily) plus MTX (20 mg/week) with monotherapy of TOF or MTX over 12 months in MTX-naïve patients with early RA.29 At 12 months, the TOF plus MTX group reached higher improvements in disease activity (DAS28-4 ESR [Disease Activity Score in 28 joints with 4 variables including erythrocyte sedimentation rate] <3.2) than either of the monotherapy groups receiving only TOF or MTX (58.8% vs. 30.6% vs. 18.9%, p<0.001); the combination group also experienced higher remission rates (DAS28-4 ESR <2.6: 35.3%, 19.4%, 13.5%; p<0.05). Finally, radiographic
changes (mTSS) were smaller for the combination group than for monotherapy with either TOF or MTX (-0.15, 0.85, 0.71; p<0.05). Overall, the SOE was insufficient for comparisons of MTX plus tsDMARD with either MTX or tsDMARD.

**Biologics**

**TNF Biologics**

**TNF Biologic Versus csDMARD Monotherapy**

Thirteen RCTs compared a TNF biologic with csDMARD monotherapy. Nearly all of these trials reported baseline similarity of patient characteristics between treatment arms, with the exception of one trial\(^{34}\) in which differences existed in terms of age, physical functional capacity, and Sharp joint space narrowing score. These differences contributed only partially to an elevated ROB rating.\(^{34}\) These trials examined the question of whether adding a TNF biologic improves outcomes in csDMARD users. TNF biologics examined included all TNF biologics except GOL — ADA, CZP, ETN, and IFX. Overall, the TNF biologics (ADA, CZP, ETN, and IFX) plus MTX have smaller radiographic changes and higher remission rates than MTX monotherapy (low SOE).

*Adalimumab.* Five RCTs, one of which was previously described in the csDMARDs versus TNF biologics section, examined the combination of ADA (40 mg biweekly) plus MTX (ranging from 8 to 20 mg/week) with MTX monotherapy over 26 weeks to 2 years.\(^{13, 15, 34-37, 103, 114-119, 150-152, 160-163}\) Results were mixed: four trials showed improvements in disease activity, and five trials showed smaller radiographic changes for the combination of ADA plus MTX; two trials showed no significant differences but trended in favor of combination therapy. One trial did not report any data about radiographic progression.\(^{16}\) The trials showing differences were conducted over a shorter period (26 weeks), whereas the longer trials did not. NWMA found higher ACR50 responses and less radiographic progression for ADA plus MTX combination therapy than for MTX (RR, 1.35; 95% CI, 1.15 to 1.59, and SMD, -0.99; 95% CI, -1.17 to -0.81, respectively) (Figure 9 for ACR50 and Figure 10 for radiographic joint damage).

Overall, the SOE for comparisons of ADA plus MTX with MTX monotherapy was low for remission and insufficient for disease activity and radiographic changes.
Figure 9. Forest plot for network meta-analysis of ACR50 response rates: Comparison of TNF plus MTX with MTX only

- Adalimumab+MTX vs. MTX: 1.35 (1.15, 1.59)
- Certolizumab+MTX vs. MTX: 1.20 (1.04, 1.38)
- Etanercept+MTX vs. MTX: 1.49 (1.27, 1.74)
- Infliximab+MTX vs. MTX: 1.57 (1.30, 1.88)

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

Figure 10. Forest plot for network meta-analysis of change from baseline in radiographic joint damage score: Comparison of TNF combined therapies with MTX only

- Adalimumab+MTX vs. MTX: -0.99 (-1.17, -0.81)
- Certolizumab+MTX vs. MTX: -0.38 (-0.53, -0.23)
- Etanercept+MTX vs. MTX: -0.81 (-0.98, -0.63)
- Infliximab+MTX vs. MTX: -0.42 (-0.58, -0.27)

SMD (95% CI)
The HIT HARD trial (n=387) was a 48-week trial of combination ADA (40 mg biweekly) plus MTX (15 mg/week) compared with ADA or MTX monotherapy in MTX-naïve patients in private rheumatology practices, hospitals, and university departments throughout Germany.34 ADA was given 40 mg subcutaneously every other week over 24 weeks. Although patients on combination therapy had significant reductions in disease activity (DAS28) at week 24, the differences in clinical outcomes were not significant at week 48 (3.2 vs. 3.4, p=0.4).

The HOPEFUL 1 trial randomized 334 MTX-naïve Japanese patients with early RA to ADA (40 mg biweekly) plus MTX (6 to 8 mg/week) to MTX monotherapy.35 After 26 weeks, remission rates (DAS28<2.6) were significantly higher for combination therapy than with MTX only (31% vs. 14.7%, p<0.001).

The largest trial, OPTIMA,37, 151, 152 was a phase 4 multinational trial that randomized 1,032 early RA patients that were MTX naïve to ADA (40 mg biweekly) plus MTX (7.5 to 20 mg/week) or MTX for 26 weeks (period 1). After period 1 (26 weeks), patients receiving combination ADA plus MTX had significantly higher ACR50 response rates (52% vs. 34%, p<0.001) and significantly lower mean Sharp/van der Heijde Method for Scoring Radiographs (SHS) radiographic changes (0.15 vs. 0.96, p<0.001).

The PREMIER study,15 previously described above in the csDMARDs vs. Biologics section (N=799) compared MTX (20 mg/week) plus the TNF biologic ADA (40 mg biweekly) with either drug alone in MTX-naïve patients with early aggressive RA. Significantly more patients on MTX plus ADA achieved an ACR50 response than did patients receiving monotherapy with either MTX or ADA (59% vs. 43% vs. 37%, p<0.001) at 2 years. Patients in the ADA plus MTX group had also higher remission rates (49% vs. 25% vs. 25%, p<0.001). Additionally, the combination therapy group had lower radiographic progression (modified Sharp/van der Heijde score [mTSS]: 1.9 vs. 5.5 vs. 10.4; p<0.001). During the 10-year open-label extension,118 patients taking ADA plus MTX had significantly less radiographic progression than those on monotherapy, but results were limited by a 34 percent overall attrition rate.

The PROWD study,16 rated high ROB, also found similar improved disease activity with ADA plus MTX combination therapy in 148 MTX-naïve patients but no significant differences in ACR50 response rates and remission at 56 weeks.

Certolizumab pegol. Two RCTs examined the combination of CZP plus MTX versus MTX monotherapy in MTX-naïve patients.13, 38 The C-OPERA trial (N=316), conducted in Japan,13, 153 randomized patients with early RA with poor prognostic factors (high anti-CCP antibody, positive RF, or bony erosions) to CZP, 400 mg biweekly for 4 weeks, then 200 mg biweekly, plus MTX (up to 20 mg/week) or to MTX only. ROB was medium at 24 weeks but high at 52 weeks and 2 years because of high attrition. At 24 weeks, patients in the CZP plus MTX group had significantly higher DAS28 ESR remission rates (52.8% vs. 30.6%, p<0.001) and significantly lower radiographic progression (modified SHS mean change 0.26 vs. 0.88, p=0.003). Similar findings were noted at 2 years.

The second trial, C-EARLY, a 52-week multinational trial38, 39 (n=879) of patients also with poor prognostic factors found significantly higher ACR50 response for patients on CZP (400 mg biweekly) plus MTX (up to 25 mg/week) (61.8% vs. 52.6%, p=0.023) and significantly higher DAS28-ESR remission (42.6% vs. 26.8%, p<0.001) than MTX monotherapy. Additionally, the CZP plus MTX group had a significantly higher proportion of patients with no radiographic progression by mTSS from baseline (70.3% vs. 49.7%, p<0.001).
In the NWMA, higher ACR50 response rates and less radiographic progression were also noted for CZP plus MTX combination therapy than MTX monotherapy (RR, 1.20; 95% CI, 1.04 to 1.38, and SMD, -0.38; 95% CI, -0.53 to -0.23, respectively) (Figure 9 for ACR50 and Figure 10 for radiographic joint damage).

Overall the SOE for comparisons of CZP plus MTX with MTX monotherapy was low for disease activity, remission, and radiographic changes.

*Etanercept.* Three trials compared ETN (25 mg twice weekly or 50 mg weekly) with MTX in MTX-naïve patients.12, 14, 113 The COMET trial included 542 patients with early RA over 2 years.12, 108, 109, 154-156 Patients were randomized into four groups: (1) ETN plus MTX for 2 years (ETN-MTX/ETN-MTX), (2) ETN plus MTX for year 1 followed by ETN alone in year 2 (ETN-MTX/ETN), (3) MTX for year 1 followed by ETN plus MTX in year 2 (MTX/ETN-MTX), or (4) MTX for 2 years (MTX/MTX). Patients in the ETN plus MTX group had a significantly higher ACR50 response than MTX monotherapy at 52 weeks (70.7% vs. 49.0%, p<0.001). Remission was also significantly higher in the ETN plus MTX group (DAS remission <2.6; 51.3% vs. 27.8%, p<0.0001). After 2 years, remission remained higher for patients in the ETN-MTX/ETN-MTX group compared with the MTX/MTX group (57.0% vs. 35.0%, p=0.002).

The Enbrel Early RA (ERA) trial found no significant difference in ACR20 response rates (65.0% vs. 72.0%, p=0.16) or radiographic changes at the primary outcome of 12 months, but the 1-year open-label extension found higher ACR20 response rates for ETN than for MTX (72.0% vs. 59.0%; p=0.005).14, 36, 110-112

The third trial113 did not find any significant differences in DAS28 between groups (3.2 vs. 3.1, p=0.53) but was of shorter duration (24 weeks) and smaller sample size (n=26).

Overall, the SOE for comparisons of ETN plus MTX with MTX monotherapy was moderate for disease activity and radiographic changes and low for remission.

In the NWMA, higher ACR50 response rates and less radiographic progression were also noted for ETN plus MTX combination therapy than MTX monotherapy (RR, 1.49; 95% CI, 1.27 to 1.74, and SMD, -0.81; 95% CI, -0.98 to -0.63, respectively) (Figure 9 for ACR50 and Figure 10 for radiographic joint damage).

*Infliximab.* Three trials examined the combination of IFX with MTX compared with monotherapy in MTX-naïve patients.17, 18, 41 The ASPIRE trial (n=1,049) compared the efficacy of initiating two different combinations of IFX (3 mg/kg or 6 mg/kg) and MTX or MTX (20 mg/week) monotherapy over 54 weeks.17, 106, 107 At 54 weeks, ACR response proportions were significantly improved for both IFX plus MTX combination therapy groups compared with MTX monotherapy (ACR50: 45.6% vs. 50.4% vs. 31.1%, p<0.001 for both IFX comparisons with MTX). Patients treated with IFX plus MTX also had higher rates of remission (DAS28 ESR <2.6; 21.3% for IFX combination therapy groups vs. 12.3%, p<0.001)106 and less radiographic progression (modified SHS change: 0.4 to 0.5 for IFX combination therapy groups, 3.7, p<0.001).17

The smaller second trial (n=20) found significantly improved ACR50 responses at 54 weeks (IFX plus MTX: 78%, MTX: 40%, p<0.05) but no significant differences in radiographic progression.41 After 54 weeks, corticosteroids were permitted as clinically required. However, at 2 years, there were no significant differences in ACR50 response rates or radiographic changes (SHS scores).

The third trial, also small (n=44) and previously described in the High-Dose Corticosteroids section, found a trend in greater improvement for IFX plus MTX compared with MTX
monotherapy in ACR20, 50, or 70, but it was not significant at 1 year between groups (results reported in graph only).18

In the NWMA, IFX plus MTX combination therapy also led to higher ACR50 response rates and less radiographic progression than MTX monotherapy (RR, 1.57; 95% CI, 1.30 to 1.88, and SMD, -0.42; 95% CI, -0.58 to -0.27, respectively) (Figure 9 for ACR50 and Figure 10 for radiographic joint damage).

Overall, the SOE for comparisons of IFX plus MTX with MTX monotherapy was low for remission and insufficient for disease activity and radiographic changes.

**TNF Biologic Versus csDMARD Combination Therapy**

One trial with ADA9 and two trials with IFX10,40 examined the role of TNF biologics compared with that of csDMARD combinations. Overall, results were mixed.

*Adalimumab.* The IMPROVED trial (N=161) was a multicenter randomized single-blind trial comparing a combination of ADA (40 mg biweekly) with MTX (25 mg/week), HCQ (400 mg/day), SSZ (2 g/day), and PRED (7.5 mg/day) plus MTX (25 mg/week) in patients who were inadequate responders to MTX.9,120,158 Initially, all patients were treated with MTX (25 mg/week) and a tapered high dose of PRED from 60 mg to 7.5 mg/day. Patients who were not in early remission (DAS 1.6 or higher) were randomized into the two treatment groups. After 2 years, no significant differences were observed for disease activity (DAS mean change: 2.0 vs. 1.9, p=0.45), remission (DAS <1.6: 26.5% vs. 30.8%, p=0.76), or radiographic progression (mTSS progression ≥0.5: 10.8% vs. 6.4%, p=0.31). Overall, the SOE for comparisons of ADA plus MTX with csDMARD combination therapy is insufficient for disease activity, remission, and radiographic changes.

*Infliximab.* The SWEFOT trial10,121-126 was a multicenter randomized trial (n=258) in Sweden comparing IFX (3 mg/kg) plus MTX with MTX (20 mg/week) plus SSZ (2 g/day) plus HCQ (400 mg/day) over 1 year in patients who were inadequate responders to MTX. Initially, 487 patients were enrolled and placed on MTX for 3 to 4 months; those who did not achieve low disease activity were randomized into the above therapies. After 1 year, the IFX plus MTX combination group had significantly higher ACR50 response rates (25.0% vs. 14.6%, p=0.042). However, in a 2-year followup study of MTX naïve patients,122 ACR50 response rates were not significantly different between groups. The 2-year followup results from the NEO-RACo trial comparing IFX plus the FIN-RACo regimen of MTX (25 mg/week) plus SSZ (1 to 2 g/d) plus HCQ (35 mg/kg/week) plus PRED (7.5 mg/day) with the FIN-RACo regimen no significant differences in ACR50, remission (61% vs. 60%, p=0.93) or radiographic progression (SHS mean: 5.3 vs., p=0.54) at 5-year followup.40,127,128 Overall, the SOE for comparisons of IFX plus MTX with csDMARD combination therapy is low for disease activity.

**Non-TNF Biologics**

**Non-TNF Biologic Plus MTX Versus Either Non-TNF Biologic or MTX**

*Abatacept.* The AGREE trial was a multinational trial of early RA patients (98% MTX naïve) with poor prognostic factors (n=509) that compared the combination ABA (10 mg/kg days 1, 15, and 29 and then every 4 weeks) plus MTX (7.5 mg/week) with MTX only over 2 years.31,129-131 The first year was a double-blind trial; in year 2, patients in the combination therapy (ABA plus MTX) continued treatment and ABA was initiated in the MTX-only group. After 1 year, the ABA plus MTX group had significantly higher ACR50 response than the MTX-only group (57.4% vs. 42.3%, p<0.001). The ABA plus MTX group also had significantly higher remission
rates (41.4% vs. 23.3%, p<0.001) and less mean radiographic changes (Genant-modified Sharp score 0.63 vs. 1.06, p=0.040). Less radiographic progression was noted at 2 years for the original ABA plus MTX group compared with progression for the original MTX-only group.\textsuperscript{130}

The multinational AVERT study (n=351), previously described in the csDMARDs versus non-TNF biologics section, also compared the combination of ABA (125 mg/week) plus MTX (7.5 mg/week) with ABA monotherapy and also MTX monotherapy (prior MTX use not reported).\textsuperscript{7} Overall, the non-TNF biologic ABA plus MTX had smaller radiographic changes (low SOE) and higher remission rates (moderate SOE) than MTX monotherapy.

The NWMA found significant differences in ACR50 response when comparing ABA plus MTX with MTX monotherapy (RR, 1.34; 95% CI, 1.16 to 1.54), consistent with the results from the AGREE and AVERT trials (Figure 11). The combination of ABA plus MTX had numerically less radiographic progression than MTX monotherapy, but the difference was not significant (SMD, -0.09; 95% CI, -0.26 to 0.09) (Figure 12).

**Figure 11. Forest plot for network meta-analysis of ACR50 response rates: Non-TNF biologic plus MTX versus MTX**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. MTX</td>
<td>1.34 (1.16, 1.54)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. MTX</td>
<td>1.29 (1.13, 1.47)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

**Figure 12. Forest plot for change from baseline in radiographic joint damage score: Non-TNF biologic plus MTX versus MTX**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. MTX</td>
<td>-0.09 (-0.26, 0.09)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. MTX</td>
<td>-0.26 (-0.40, -0.12)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; SMD = standardized mean difference (mean difference divided by standard deviation); TNF = tumor necrosis factor; vs. = versus.
**Rituximab.** The IMAGE trial\(^{30,132,133}\) (n=755) randomized MTX-naïve patients to RIT (1 g days 1 and 15) plus MTX (7.5 mg-20 mg/week) combination therapy, RIT (500 mg days 1 and 15) plus MTX (7.5 mg to 20 mg/week) combination therapy, and MTX monotherapy over 52 weeks. Both RIT plus MTX groups and the RIT monotherapy group had significantly improved disease activity (DAS28: -3.21 vs. -3.05 vs. -2.06, p<0.001) and remission rates (31% vs. 25% vs. 13%, p<0.001) and less radiographic change (0.36 vs. 0.65 vs. 1.08, p<0.001 compared with MTX monotherapy). Overall, the non-TNF biologic RIT plus MTX had smaller radiographic changes (moderate SOE) and higher remission rates (moderate SOE) than MTX monotherapy.

In the NWMA, TCZ plus MTX showed higher ACR50 response rates and less radiographic progression than MTX monotherapy (RR, 1.29; 95% CI, 1.13 to 1.47, and SMD, -0.26; 95% CI, -0.40 to -0.12, respectively) (Figure 11 for ACR50 and Figure 12 for radiographic joint damage). There were no NWMA comparisons with RIT or SAR.

**Tocilizumab.** Two RCTs, the FUNCTION trial\(^{32}\) (N=1,162) and the U-Act-Early trial\(^{33}\) (N=317), both previously described in the csDMARD versus non-TNF biologic section, assessed differences in efficacy between a TCZ plus MTX combination and either MTX or TCZ monotherapy in MTX-naïve populations. In both trials, the non-TNF biologic TCZ plus MTX led to smaller radiographic changes (moderate SOE) and higher remission rates (low SOE) than MTX monotherapy after 1 to 2 years.

**Biologic Head to Head: TNF Versus Non-TNF**

The ORBIT trial, an open-label noninferiority RCT (n=329), compared the non-TNF, RIT (1 g days 1 and 15) with TNF, ADA (40 mg biweekly), or ETN (50 mg/week) over 1 year.\(^8\) Patients had a prior inadequate response to at least two csDMARDs. Despite attempting two treatments, the mean disease duration was 6.7 to 8.0 months. No significant differences were found for disease activity (DAS28 ESR mean change: -2.6 vs. -2.4, p=0.24) or remission (DAS28 remission: 23% vs. 21%, p=NR). Radiographic progression was not reported. Overall, the SOE for the comparison of TNF with non-TNF therapies was insufficient.

In the NWMA below (Figure 13 for ACR50 and Figure 14 for radiographic joint damage), TNF therapy (monotherapy or with MTX) is compared with non-TNF therapy (monotherapy or with MTX). No comparisons were significant, except for a lower ACR50 response rate for ADA compared with TCZ (RR, 0.75; 95% CI, 0.58 to 0.95). Less radiographic progression was noted with ADA plus MTX (SMD, -0.90; 95% CI, -1.15 to -0.65) and CZP plus MTX (SMD, -0.29; 95% CI, -0.53 to -0.06) than ABA plus MTX. Less radiographic progression was also noted with ADA plus MTX than TCZ plus MTX (SMD, -0.73; 95% CI, -0.96 to -0.50).
Figure 13. Forest plot for network meta-analysis of ACR50 response rates: TNF biologic versus non-TNF biologic

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

Figure 14. Forest plot for change from baseline in radiographic joint damage score: TNF biologic versus non-TNF biologic

95% CI = 95% confidence interval; MTX = methotrexate; SMD = standardized mean difference (mean difference divided by standard deviation); TNF = tumor necrosis factor; vs. = versus.
No direct evidence was available for comparisons of TNF biologics with TNF biologics. The SOE for all indirect estimates was low (downgrading for indirectness and imprecision in all cases). NWMA of ACR50 response rates found no significant differences in comparisons with ADA plus MTX versus CZP plus MTX, ETN plus MTX, or IFX plus MTX. IFX plus MTX had higher ACR50 response rates than CZP plus MTX, but the confidence interval was large (RR, 1.30; 95% CI, 1.04 to 1.64) (Figure 15). Radiographic progression was less for ADA plus MTX compared with IFX plus MTX (SMD, 0.57; 95% CI, 0.33 to 0.80) and CZP plus MTX (SMD 0.61; 95% CI, 0.37 to 0.84). ADA monotherapy also had less radiographic progression than ETN monotherapy (SMD, -0.49; 95% CI, -0.75 to -0.23). Radiographic progression was less for ETN plus MTX compared with CZP plus MTX (SMD, -0.42; 95% CI, -0.66 to -0.19) and IFX plus MTX (SMD, 0.38; 95% CI, 0.15 to 0.62) (Figure 16).

**Figure 15. Forest plot for network meta-analysis of ACR50 response rates: TNF biologic versus TNF biologic**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>0.89 (0.72, 1.10)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>1.10 (0.88, 1.38)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>1.16 (0.91, 1.48)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Certolizumab+MTX</td>
<td>1.24 (1.00, 1.53)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Certolizumab+MTX</td>
<td>1.30 (1.04, 1.64)</td>
</tr>
<tr>
<td>Adalimumab vs. Etanercept</td>
<td>0.94 (0.72, 1.24)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Etanercept+MTX</td>
<td>1.05 (0.83, 1.34)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
Figure 16. Forest plot for change from baseline in radiographic joint damage score: TNF biologic versus TNF biologic

95% CI = 95% confidence interval; MTX = methotrexate; SMD = standardized mean difference (mean difference divided by standard deviation); TNF = tumor necrosis factor; vs. = versus.

Non-TNF Versus Non-TNF

No direct evidence was available for comparisons of non-TNF biologics with non-TNF biologics. The SOE for all indirect estimates was low (downgrading for indirectness and imprecision in all cases). In NWMA of ACR50 response and radiographic progression, comparisons of TCZ (with or without MTX) versus ABA (with or without MTX) found no significant differences between groups (low SOE) (Figure 17 and Figure 18, respectively).

Figure 17. Forest plot for network meta-analysis of ACR50 response rates: Non-TNF biologic versus non-TNF biologic

95% CI = 95% confidence interval; ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; v TNF = tumor necrosis factor; s. = versus.
Combinations and Therapy Strategies

With respect to combination therapy, long-term studies show no differences in remission rates between initial combination versus step-up therapies (moderate SOE). The BeSt study randomized 508 MTX-naïve patients with early RA to one of four groups: (1) sequential DMARD, starting with MTX (15 mg/week); (2) stepped-up combination therapy with MTX (15-30 mg/week) followed by SSZ (2 g/day), HCQ, and PRED; (3) initial combination therapy of MTX, SSZ, and tapered high-dose PRED (60 mg/day to 7.5 mg/day in 7 weeks); and (4) initial combination therapy with MTX (25-30 mg/week) and IFX (3 mg/kg every 8 weeks; doses titrated up to 10 mg/kg dependent on DAS >2.4).79-91 The design called for frequent changes in drug strategy; therapeutic strategies were adjusted every 3 months when the DAS was greater than 2.4. At 12 months, higher proportions in group 3 (MTX, SSZ, PRED) and group 4 (MTX and IFX) reached a DAS of 2.4 or less (group 1: 53%; group 2: 64%; group 3: 71%; and group 4: 74%, p=0.004 for group 1 vs. group 3, p=0.001 for group 1 vs. group 4: p=NS for other comparisons).79 The median increase in total SHS radiographic scores was 2.0, 3.5, 1.0, and 0.5 in groups 1 through 4 (p<0.001),79 suggesting that initial combination therapies resulted in less radiographic damage. At 4 years, remission rates were similar among the groups (DAS <1.6: 50%, 41%, 38%, 42%, p=0.40).86 Similarly, there were no significant differences among the groups in remission at 10 years (51.0%, 49.0%, 53.0%, 53.0%, p=0.94). There were also no significant differences in joint damage at 10 years (mTSS: 11.0, 8.0, 8.0, 6.0, p=0.15).

The GUEPARD study92 first randomized MTX-naïve patients to 3 months of ADA plus MTX or MTX monotherapy. In patients who at 3 months did not respond to an initial strategy, investigators examined whether disease activity-driven treatment with TNF inhibitors was equally effective in controlling clinical symptoms and structural damage in both groups. At 3 months, there was an initial numerical improvement in ACR50 response (66% vs. 27%, p=NR), but there were no differences at 1 year between groups. Similarly, there were no differences in radiographic changes between groups. We rated this study high ROB after 12 weeks because of the risk of contamination bias given that patients could be switched to different dosing and treatment regimens when low disease activity was achieved at 12 weeks and beyond (both groups received the same treatments).
Similarly, the OPERA trial\textsuperscript{36, 160-163} of 180 early RA patients in Danish hospital-based clinics using a treat-to-target protocol found no significant differences in disease activity or remission between combination therapy (ADA plus MTX) and monotherapy (MTX) (DAS28 CRP [Disease Activity Score based on C-Reactive Protein]<2.6 remission: 66% vs. 69%, p=0.79).

The TEAR study\textsuperscript{20, 159} randomized MTX-naïve patients (n=755) to four treatment arms: (1) immediate treatment with MTX plus ETN; (2) immediate treatment with MTX plus SSZ plus HCQ (triple therapy); (3) step-up from MTX to MTX plus ETN when DAS28-ESR (Disease Activity Score 28 using erythrocyte sedimentation rate) was 3.2 or higher at week 24; and (4) step-up from MTX to triple therapy when DAS28-ESR was 3.2 or higher at week 24. The four treatment groups did not differ significantly in DAS28-ESR between week 48 and week 102 (reported in figure only, p=0.48). Similarly, radiographic score changes (mTSS) did not differ significantly between step-up therapy and immediate therapy. Radiographic progression was significantly lower among patients randomized to MTX plus ETN than among those receiving triple therapy (0.64 vs. 1.69, p=0.047). We rated this trial as high ROB because overall discontinuation rates were high (up to 42 percent).

**KQ 2: Comparative Benefits of Drug Therapies for Patients With Early RA in Relation to Patient-Reported Symptoms, Functional Capacity, or Quality of Life**

To address this KQ, we had a total of 41 studies (40 RCTs and 1 observational study). Details of individual studies are documented in the Evidence Table in Appendix C; some information about the specific investigations that had also addressed KQ 1 can be found in the “Characteristics of Included Studies” section above.
Table 9 presents data on all these investigations for the three main outcomes of concern for KQ 2: patient-reported symptoms, functional capacity (sometimes denoted as function or physical function), and quality of life (typically health-related quality of life, or HRQOL). Functional capacity was the most commonly measured outcome. HAQ-DI was the most common outcome measure reported for physical function. The accepted minimally clinically important difference (MCID) for HAQ-DI in RA is a change of 0.22-0.25. HRQOL was sometimes assessed, and 36-item Short Form Health Survey (SF-36) Physical Component Score (PCS) and SF-36 Mental Component Score (MCS) were the most common outcome measures reported for HRQOL. The accepted MCID for the SF-36 PCS in RA is 4.4, and for the SF-36 MCS, it is 3.1. Patient-reported symptoms were only rarely reported. Appendix F provides more information about the scales and their meanings.

Key Points

- Conclusions below are based on early RA studies including patients with moderate to high disease activity, and the majority were MTX naive.
- Evidence was insufficient to determine the impact of corticosteroids plus csDMARDs versus csDMARD monotherapy on functional capacity or health-related quality of life (HRQOL).
- Combinations of TNF biologics plus MTX produced statistically significantly greater improvements in functional capacity than MTX alone. The differences in HAQ-DI exceeded the minimally clinically important difference in most studies. This finding applied to the following TNF biologics: ADA (difference of HAQ change -0.1 to -0.3 over 24 weeks to 2 years) (moderate SOE), CZP (difference of HAQ change not consistently reported, but in favor of combination therapy, over 30 weeks to 1 year) (low SOE), and IFX (difference of HAQ change not consistently reported, but in favor of combination therapy, over 30 weeks to 1 year) (low SOE). Evidence was inconclusive for the TNF biologic ETN (low SOE). Evidence was insufficient to determine the impact on HRQOL of adding TNF biologics to MTX therapy.
- The TNF biologic IFX plus a combination of csDMARDs (triple therapies—MTX, SSZ, HCQ, plus prednisone [PRED]) did not differ significantly from the same combination of csDMARDs alone in their impact on functional capacity (low SOE). Evidence was insufficient to determine whether ADA plus MTX or IFX plus MTX differed from csDMARD triple therapy in their effects on functional capacity.
- Combination of RIT (non-TNF biologic) plus MTX produced statistically significantly greater improvements in functional capacity than MTX alone (HAQ decrease >0.22: 88% and 87% vs. 77%, p<0.05) (moderate SOE).
- Evidence was insufficient to evaluate any differences between one biologic and another biologic for their impact on either functional capacity or HRQOL.
- Combination strategies using multiple csDMARDs or csDMARD plus TNF biologics compared with sequential or step-up therapies did not differ significantly in terms of functional capacity (low SOE). Evidence was insufficient to determine the impact of these strategies on HRQOL.
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr</th>
<th>Study Design</th>
<th>N</th>
<th>Duration</th>
<th>Comparison (Dose)</th>
<th>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CAMERA-II, 2012&lt;sup&gt;24&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=239</td>
<td>2 yrs</td>
<td>PRED (10 mg/day) + MTX (10 mg/wk) vs. MTX (10 mg/wk)</td>
<td>Higher mean HAQ score in MTX vs. MTX + PRED at 2 yrs (0.7 vs. 0.5), Mean difference (95% CI): -0.18 (-0.34 to -0.02) (p=0.027). Similar statistically significant differences were found at 3, 6, 12, and 18 months.</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CARDERA, 2007&lt;sup&gt;77&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=467</td>
<td>2 yrs</td>
<td>PNL (60 mg/day tapered over 34 wks) + MTX (7.5-15 mg/wk) vs. MTX</td>
<td>At 2 yrs, no difference in HAQ mean change in MTX + PNL vs. MTX (-0.28 vs. -0.29, p=NR) Mean increase in SF-36 PCS was 5.8. No difference in the SF-36 PCS mean change between MTX and MTX + PNL (p=NR). No difference in SF-36 MCS or EQ-5D between groups.</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>Montecucco et al., 2012&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Open label RCT</td>
<td>N=220</td>
<td>12 months</td>
<td>PRED (12.5 mg/day for 2 weeks then taper to 6.25 mg/day) + MTX (10-25 mg/week) vs. MTX (10-25 mg/week)</td>
<td>More improvement in patient-reported pain (VAS, mean change) in the PRED + MTX group than in the MTX group at 4 and 12 months, but not 6 or 9 months</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CareRA, 2015&lt;sup&gt;95, 98&lt;/sup&gt;, 2017&lt;sup&gt;99&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=379</td>
<td>2 yrs</td>
<td>High-risk patients: 1: MTX (15 mg/wk) + SSZ (2 g/day) + PRED (60 mg/day tapered to 7.5 mg/day) vs. 2: MTX + PRED (30 mg tapered to 5 mg/day) vs. 3: MTX + LEF (10 mg/day) + PRED (30 mg tapered to 5 mg/day) vs. Low-risk patients: 4: MTX 15 mg/wk vs. 5: MTX + PRED (30 mg tapered to 5 mg/day)</td>
<td>No differences in functional capacity among the groups at 16 weeks and 54 weeks as measured by clinically meaningful change in HAQ change (p= NS). Fewer patients had a HAQ score of 0 in the MTX-TSU group (23.4%) than in the COBRA Slim group (51.2%) (p=0.006).</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>BARFOT #2, 2005&lt;sup&gt;78&lt;/sup&gt;, 2009&lt;sup&gt;97&lt;/sup&gt;, 2014&lt;sup&gt;138, 140&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=259</td>
<td>2 yrs</td>
<td>PNL 7.5 mg/day + DMARD (SSZ 2 g/day or MTX 10 mg/wk) vs. DMARD (SSZ 2 g/day or MTX 10 mg/wk)</td>
<td>Significant improvement in physical function as measured by mean decrease in HAQ from baseline between the PNL + csDMARD group compared with the csDMARD group at all time points including 3, 6, 12, 18 months and 2 yrs (p=0.003). Significant difference between groups still present at 4 yrs (p=0.034). Patients in remission at 2 yrs had significantly lower HAQ scores at both 2 and 4 yrs.</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Rating</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------</td>
<td>---------------------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>High-dose corticosteroids</td>
<td>Durez et al., 2007&lt;sup&gt;11, 2&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>18</td>
<td>1 yr</td>
<td>IFX 3 mg/kg 0.2,6 and every 8 wks + MTX (7.5-20 mg/wk) vs. MTX + Methyl-PNL (1 g at 0.2,6 and every 8 wks) vs. MTX</td>
</tr>
<tr>
<td>High-dose corticosteroids</td>
<td>IDEA, 2014&lt;sup&gt;96&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>96</td>
<td>26 weeks</td>
<td>IFX (3 mg/kg at wks 0, 2, 6, 14, 22) + MTX (10 to 20 mg/wk) vs. Methyl-PNL (250 mg single dose) vs. MTX</td>
</tr>
<tr>
<td>csDMARD Monotherapy vs. csDMARD Monotherapy</td>
<td>BARFOT #1, 2003&lt;sup&gt;37&lt;/sup&gt;</td>
<td>High</td>
<td>RCT</td>
<td>245</td>
<td>2 yrs</td>
<td>PNL (7.5-15 mg/day for 1-3 months) + MTX (5-15 mg/wk) vs. SSZ (2-3g/day) + PNL (up to 10 mg/day)</td>
</tr>
<tr>
<td>csDMARD Monotherapy vs. csDMARD Monotherapy</td>
<td>NOR-DMARD, 2012&lt;sup&gt;28&lt;/sup&gt;</td>
<td>High</td>
<td>Observational</td>
<td>1,102</td>
<td>3 yrs</td>
<td>SSZ (2 g/day) vs. MTX (10 mg-15 mg/wk)</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Dougados et al., 1999&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>209</td>
<td>1 yr (5-yr followup)</td>
<td>SSZ (2-3g/day) + MTX (7.5 to 15 mg/wk) vs. SSZ vs. MTX</td>
</tr>
<tr>
<td></td>
<td>Maillefert et al., 2003&lt;sup&gt;184&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>105</td>
<td>1 yr</td>
<td>SSZ (1-3 g/day) + MTX (7.5-15 mg/wk) vs. MTX vs. SSZ</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Rating</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>---------------------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>COBRA, 1997, 2002, 2009</td>
<td>Medium</td>
<td>RCT</td>
<td>155</td>
<td>5 yrs</td>
<td>PNL (60 mg tapered over 28 wks) + MTX (7.5 mg/wk stopped after 40 wks) + SSZ (2 g/day) vs. SSZ</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>COBRA-Light, 2014</td>
<td>Medium</td>
<td>RCT</td>
<td>164</td>
<td>1 yr</td>
<td>PNL (60 mg tapered over 28 wks) + MTX (7.5 mg/wk) + SSZ (2,000 mg/day) (“COBRA”) vs. PNL (30 mg tapered over 28 wks), MTX (7.5 mg to 25 mg/wk) “COBRA Light”)</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>FIN-RACo, 1999, 2004, 2010, 2013</td>
<td>Medium</td>
<td>RCT</td>
<td>199</td>
<td>2 yrs</td>
<td>MTX (7.5-10 mg/wk) + HCQ (300 mg/day) + SSZ (2 g/day) + PNL (5-10 mg/day) vs. DMARD (SSZ could be changed to MTX if adverse event or lack of response)</td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>tREACH, 2013, 2014, 2016</td>
<td>Medium</td>
<td>RCT</td>
<td>515</td>
<td>1 yr</td>
<td>MTX (25 mg/wk) + SSZ (2 g/day) + HCQ (400 mg/day) + glucocorticoid IM vs. MTX + SSZ + HCQ + glucocorticoid oral taper (15 mg/day tapers off at 10 wks) vs. MTX + glucocorticoid oral taper</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr Risk of Bias Rating</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD + TNF Biologic vs. TNF Biologic</td>
<td>PREMIER, 2006,15 2008,100 2010,115,149 2012,116 2013,117 2014,118 2015119 c Medium</td>
<td>RCT N=799 2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (20 mg/wk) vs. ADA vs. MTX</td>
<td>At 3 months and 6 months, no significant differences in function or HRQOL between groups. At 1 yr, HAQ-DI mean change was greater in the ADA + MTX group than in both the ADA group (p=0.0002) and the MTX group (p=0.0003). At 76 weeks, no significant difference in SF-36 scales or pain. Function improved significantly more in the ADA + MTX group than in the MTX group (HAQ-DI mean change: -1 vs. -0.9, p&lt;0.05; HAQ-DI response, p=NS). Significantly more patients in the ADA + MTX group had a HAQ-DI score of 0 than in either monotherapy group (33% vs. 19% vs. 19%, p&lt;0.001). SF-36 PCS improved more in ADA + MTX group than in MTX group (p&lt;0.0001); no difference in MCS. SF36 MCS improved more in the ADA group than the MTX group (p=0.015). Patient-reported pain (VAS, mean) was lower in the ADA + MTX group than the ADA group (p&lt;0.0001). No difference between the ADA and MTX groups. More days of employment and fewer missed work days in the ADA + MTX group than in the MTX group.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD + Non-TNF Biologic vs. csDMARD</td>
<td>AVERT, 201512 d Medium</td>
<td>RCT N=351 2 yrs</td>
<td>ABA (125 mg/wk) + MTX (7.5-15 mg/wk) vs. ABA vs. MTX</td>
<td>At 12 and 18 months: nonsignificant but higher percentages of patients in the ABA + MTX group than in the ABA group and the MTX group with HAQ-DI response (respectively by time points, 65.5% vs. 52.6% vs. 44%; 21.8% vs. 16.4% vs. 10.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD + Non-TNF Biologic vs. csDMARD</td>
<td>FUNCTION, 201612 d Medium</td>
<td>RCT N=1,162 2 yrs*</td>
<td>TCZ (4 mg/kg monthly) + MTX (20 mg/wk) vs. TCZ (8 mg/kg monthly) + MTX vs. TCZ vs. MTX</td>
<td>At 52 weeks, significantly greater improvement in mean HAD-DI scores from baseline in TCZ 8 mg + MTX group than in MTX group (p=0.0024). At 24 weeks and at 52 weeks: Significantly greater change in SF-36 PCS scores in the TCZ 8 mg/kg + MTX group than in the MTX group (p=0.0014 and p=0.0066 for both time points). No differences in SF-36 PCS scores between the TCZ 4 mg/kg + MTX group and the MTX group or between TCZ and MTX group. No differences in SF-36 MCS scores</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Note:**
- RCT: Randomized Controlled Trial
- N: Number of participants
- Duration: Duration of the study
- Comparison: Comparison between groups
- Results: Summary of patient-reported outcomes, functioning, and quality of life.
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr Risk of Bias Rating</th>
<th>Study Design N Duration</th>
<th>Comparison (Dose)</th>
<th>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</th>
</tr>
</thead>
</table>
| csDMARD + Non-TNF Biologic vs. csDMARD                                                        | U-Act-Early, 2016<sup>13</sup> Medium | RCT N=317 2 yrs         | TCZ (8 mg/kg IV monthly) + MTX 10-30 mg/wk) vs. TCZ vs. MTX | At 24 weeks, physical function differed significantly (HAQ Dutch) between TCZ + MTX group and each monotherapy group (p=0.0275)  
  At 52 weeks and 2 yrs, physical function did not differ significantly (from baseline measures) between groups  
  Significantly greater improvement in mean SF-36 PCS over time in TCZ + MTX group and TCZ monotherapy group vs. MTX monotherapy group (p=0.044 and p=0.012, respectively). No differences in SF-36 MCS over time between groups.  
  Significantly greater improvement in mean EQ-5D scores over time in TCZ + MTX group vs. MTX monotherapy group (p=0.018). No significant difference between TCZ and MTX monotherapy groups. |
| csDMARDs vs. tsDMARDs                                                                         | Conaghan et al., 2016<sup>16</sup> Medium | RCT N=108 1 yr          | TOF (20 mg/d) + MTX (10-20 mg/wk) vs. TOF vs. MTX         | At 3, 6, and 12 months, no significant differences in improvement in function (HAQ-DI) between the TOF + MTX group and either the MTX or the TOF groups |
| TNF Biologic vs. csDMARD Monotherapy                                                           | HIT HARD, 2013<sup>14</sup> Medium | RCT N=172 48 weeks      | ADA (40 mg biwkly for 24 wks) + MTX (15 mg/wk) vs. MTX    | At 24 weeks:  
  Significantly greater physical function in ADA+MTX group than in MTX group (HAQ-DI mean 0.49 vs. 0.72, p=0.0014)  
  Significantly greater SF-36 PCS (44.0 vs. 39.8, p=0.0002)  
  No difference in SF-36 MCS at 24 weeks  
  At 48 weeks: no difference between groups in function or quality of life measures |
| TNF Biologic vs. csDMARD Monotherapy                                                           | HOPEFUL 1, 2014<sup>15, 150</sup> Medium | RCT N=334 1 yr          | ADA (40 mg biwkly) + MTX (6-8 mg/wk) vs. MTX             | At 26 weeks, significantly greater improvement from baseline in physical function in ADA + MTX group than in MTX group (decrease from baseline in mean HAQ-DI score: 0.6±0.6 vs. 0.4±0.6, p<0.001)  
  At 26 weeks, significantly more patients in ADA + MTX group than in MTX group achieved normal functionality (HAQ-DI score <0.5: 60.0% vs. 36.8%, p=0.001) |
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr Risk of Bias Rating</th>
<th>Study Design N Duration</th>
<th>Comparison (Dose)</th>
<th>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>OPTIMA, 2013, 2014, 2016,152 Low</td>
<td>RCT N=1,032 78 weeks</td>
<td>ADA (40 mg biwkly) + MTX (7.5-20 mg/wk) vs. MTX</td>
<td>At week 26: Significantly greater functional improvements in ADA + MTX group than in MTX group (HAQ-DI mean score: 0.7 vs. 0.9, p&lt;0.001). Significantly greater proportion of ADA + MTX patients than MTX patients had normal function (40.0% vs. 28.0%, respectively, p&lt;0.001)</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>PREMIER, 2006, 2008, 2010, 2012, 2013, 2014, 2015,151 Medium</td>
<td>RCT N=799 2 yrs</td>
<td>ADA (40 mg biwkly) + MTX (20 mg/wk) vs. ADA vs. MTX</td>
<td>At 3 months and 6 months, no significant differences in function or HRQOL between groups. At 1 yr, HAQ-DI mean change was greater in the ADA + MTX group than in both the ADA group (p=0.0002) and the MTX group (p=0.0003). At 76 weeks, no significant difference in SF-36 scales or pain. At 2 yrs: Function improved significantly more in the ADA + MTX group than in the MTX group (HAQ-DI mean change: -1 vs. -0.9, p&lt;0.05; HAQ-DI response, p=NS). Significantly more patients in the ADA + MTX group had a HAQ-DI score of 0 than in either monotherapy group (33% vs. 19% vs. 19%, p&lt;0.001). SF-36 PCS improved more in ADA + MTX group than in MTX group (p&lt;0.001); no difference in MCS SF36 MCS improved more in the ADA group than the MTX group (p=0.015). More days of employment and fewer missed work days in the ADA + MTX group than in the MTX group</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>PROWD, 2008, 2010, 2016,151 Medium</td>
<td>RCT N=148 54 weeks</td>
<td>ADA 40 mg subcutaneous every 2 wks + MTX (7.5-25 mg/wk) vs. MTX (7.5-25 mg/wk)</td>
<td>At 16 weeks, fewer patients in the ADA + MTX group than in the MTX had job loss, although difference was statistically NS (12 [16%] vs. 20 [27.3%], p=0.092). At 56 weeks, job loss was significantly lower with ADA + MTX (-18.6%) than MTX (-39.7%, p&lt;0.005). At 56 weeks, function from baseline improved significantly in the ADA + MTX group compared with the MTX group (change in HAQ from baseline: -0.7 vs. -.04, p=0.005)</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>C-OPERA, 2016, 2017,153 Medium</td>
<td>RCT N=316 2 yrs</td>
<td>CZP (400 mg biwkly x 4 wks, then 200 mg biwkly) + MTX (8-12 mg/wk) vs. MTX</td>
<td>At 52 weeks, significantly greater improvement in HAQ-DI in the CZP + MTX group than in the MTX group. At 2 yrs, no significant difference in HAQ remission between groups (73.0% vs. 63.7%, p=0.09)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Rating</td>
<td>Study Design</td>
<td>N</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------</td>
<td>----------------------</td>
<td>--------------</td>
<td>---</td>
</tr>
</tbody>
</table>
| TNF Biologic vs. csDMARD Monotherapy | C-EARLY, 2017[38, 39] | Medium | RCT | 879 | 1 yr | CZP (400 mg biwkly X 4 wks, then 200 mg biwkly) + MTX (10-25 mg/wk) vs. MTX | At 52 weeks, significantly greater improvement in function in the CZP + MTX group than in the MTX group (HAQ-DI mean change from baseline -1.00 vs. -0.82, p<0.001)  
Significantly more patients in the CZP+MTX group than in the MTX group achieved normal function (HAQ-DI <0.5: 48.1% vs. 35.7%, p=0.002)  
More improvement in fatigue (by BRAF-MDQ) and work productivity (by WPS-RA) in the CZP + MTX group across all questions  
At all weeks preceding (12, 20, 24, 36, and 40), similar greater improvements in CZP + MTX were seen |
Significantly greater improvement in function in the ETN + MTX group than in the MTX group (HAQ, mean change: -1.02 vs. -0.72, p<0.0001)  
Significantly more patients in the ETN + MTX group than in the MTX group achieved normal function (HAQ-DI <0.5: 55% vs. 39%, p=0.0004)  
Significantly higher SF-36 PCS scores in the ETN + MTX group than in the MTX group (13.7 vs.10.7, p=0.003)  
Improvement in following work-related outcomes favoring the ETN + MTX group:  
Fewer patients had to stop working: 8.6% vs. 24% (p=0.004)  
Less absenteeism: 14.2 vs. 31.9 missed workdays |
| TNF Biologic vs. csDMARD Monotherapy | Enbrel ERA, 2000, 2003, 2005, 2006[14, 10] | Medium | RCT | 632 | 1 yr | ETN (25 mg twice wkly) vs. MTX (20 mg/wk) | At 12 months, no difference in function between groups (mean HAQ)  
In the open-label extension until 24 months, significantly more patients in the ETN group than in the MTX group achieved improvement in function (HAQ improvement >0.5 units: 37% vs. 55%, p<0.001) |
| TNF Biologic vs. csDMARD Monotherapy | Marcora et al., 2006[111] | Medium | RCT | 26 | 6 months | ETN (25 mg twice wkly) vs. MTX (7.5-20 mg/wk) | At baseline, HAQ mean was 1.9 vs. 1.2 for ETN and MTX groups, respectively  
At 12 weeks, HAQ mean was 1.2 vs. 0.6 for ETN and MTX groups, respectively  
At 24 weeks, HAQ mean was 1.0 vs. 0.6 for ETN vs. MTX groups, respectively |
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr</th>
<th>Study Design</th>
<th>N</th>
<th>Duration</th>
<th>Comparison (Dose)</th>
<th>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>ASPIRE, 2004&lt;sup&gt;17&lt;/sup&gt;, 2006&lt;sup&gt;17&lt;/sup&gt;, 2009&lt;sup&gt;106&lt;/sup&gt;, 2017&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Medium RCT</td>
<td>1,049</td>
<td>54 weeks</td>
<td>IFX (3 mg/kg/8 wks) + MTX (20 mg/wk) vs. IFX (6 mg/kg/8 wks) + MTX vs. MTX</td>
<td>At 54 weeks: significantly greater improvements in HAQ scores from baseline in both the IFX (3 mg/kg) + MTX and IFX (6 mg/kg) + MTX groups than in the MTX group (% with HAQ increase ≥0.22 units from baseline: 76%, 75.5%, 65.2%, p&lt;0.004) From 30-54 weeks: significantly greater HAQ improvements in both IFX (3 mg/kg) + MTX and IFX (6 mg/kg) + MTX groups than in the MTX group (mean decrease in HAQ scores from baseline: 0.88, 0.80, vs. 0.68, p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 54 weeks: Significantly higher SF-36 PCS in both the IFX + MTX groups than in the MTX group (11.7, 13.2, vs. 10.1, p=0.003) Significant improvements in IFX (either 3 mg/kg or 6 mg/kg) + MTX group than in the MTX group in employability (OR, 2.4, p&lt;0.001) Fewer patients were unemployable in the IFX (either 3 mg/kg or 6 mg/kg) + MTX group than in the MTX group (8% vs. 14%, p=0.05) No differences between groups in employment rate (0.5% vs. 1.3%, p&gt;0.05)</td>
</tr>
<tr>
<td></td>
<td>Quinn et al., 2005&lt;sup&gt;41&lt;/sup&gt;</td>
<td>Medium RCT</td>
<td>20</td>
<td>2 yrs</td>
<td>IFX 3 mg/kg 0, 2, 6 and every 8 wks + MTX (7.5-25 mg/wk) vs. MTX (7.5-25 mg/wk)</td>
<td>At 54 weeks, significant functional benefit (by HAQ) favoring IFX + MTX over MTX (p=0.05)</td>
</tr>
<tr>
<td></td>
<td>Durez et al., 2007&lt;sup&gt;18 b&lt;/sup&gt;</td>
<td>Medium RCT</td>
<td>44</td>
<td>1 yr</td>
<td>IFX 3 mg/kg 0,2,6 and every 8 wks + MTX (7.5-20 mg/wk) vs. MTX + Methyl-PNL (1 g at 0,2,6 and every 8 wks) vs. MTX</td>
<td>At 52 weeks, significantly greater HAQ improvements over time in IFX + MTX and methyl-PNL + MTX groups than in the MTX group (p=0.001)</td>
</tr>
<tr>
<td></td>
<td>IMPROVED, 2013&lt;sup&gt;9&lt;/sup&gt;, 2014&lt;sup&gt;158&lt;/sup&gt;, 2016&lt;sup&gt;120&lt;/sup&gt;</td>
<td>High RCT</td>
<td>161</td>
<td>2 yrs</td>
<td>ADA (40 mg biwkly) + MTX (25 mg/wk) vs. MTX + PRED (7.5 mg/day) + HCQ (400 mg/day) + SSZ (2 g/day)</td>
<td>At 4, 8, 12, and 24 months: Mean HAQ scores did not differ between groups (respectively by time points: 0.86 vs. 0.88, p=0.77; 0.74 vs. 0.81, p=0.51; 0.87 vs. 0.81, p=0.6; 0.90 vs. 0.83) SF-36 PCS and MCS did not differ by group at any time point. At 12 months, lower patient-reported pain (VAS, mean) in the ADA +MTX group.</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD</td>
<td>SWEFOT, 2009, 2012, 2013, 2015, 2016</td>
<td>RCT</td>
<td>258</td>
<td>1 yr</td>
<td>IFX (3 mg/kg at 0, 2, 6 wks then biwkly) + MTX (20 mg/wk) vs. MTX + SSZ (2 g/day) + HCQ (400 mg/day)</td>
<td>At 12 months, EQ-5D dimensions did not differ significantly between groups</td>
</tr>
<tr>
<td>Combination Therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SWEFOT, 2009, 2012, 2013, 2015, 2016</td>
<td>RCT</td>
<td>258</td>
<td>1 yr</td>
<td>IFX (3 mg/kg at 0, 2, 6 wks then biwkly) + MTX (20 mg/wk) vs. MTX + SSZ (2 g/day) + HCQ (400 mg/day)</td>
<td>At 12 months, EQ-5D dimensions did not differ significantly between groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD</td>
<td>NEO-RACO, 2013, 2014, 2015</td>
<td>RCT</td>
<td>99</td>
<td>2 yrs</td>
<td>IFX (3 mg/kg) + MTX (25 mg/wk) + SSZ (2 g/day) + HCQ (35 mg/kg/wk) + PRED (7.5 mg/day) for 26 wks vs. FIN-RACo</td>
<td>At 2 and 5 yrs, mean HAQ scores did not differ significantly between groups</td>
</tr>
<tr>
<td>Combination Therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD</td>
<td>AGREE, 2009, 2011, 2015</td>
<td>RCT</td>
<td>509</td>
<td>2 yrs</td>
<td>ABA (10 mg/kg) + MTX (7.5 mg/wk) vs. MTX</td>
<td>At 1 yr, significantly greater functional benefit in the ABA + MTX group than in the MTX group (HAQ-DI % change of &gt;0.3 units from baseline: 71.9% vs. 62.1%, p=0.024)</td>
</tr>
<tr>
<td>Monotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 1 yr, significantly greater improvement in SF-36 scales in the ABA + MTX group than in the MTX group: SF-36 MCS (8.15 vs. 6.34, p=0.046) and SF-36 PCS (11.68 vs. 9.18, p=0.005)</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD</td>
<td>AVERT, 2015</td>
<td>RCT</td>
<td>351</td>
<td>2 yrs</td>
<td>ABA (125 mg/wk) + MTX (7.5-15 mg/wk) vs. ABA vs. MTX</td>
<td>At 12 and 18 months: nonsignificant but higher percentages of patients in the ABA + MTX group than in the ABA group and the MTX group with HAQ-DI response (respectively by time points, 65.5% vs. 52.6% vs. 44%; 21.8% vs. 16.4% vs. 10.3%)</td>
</tr>
<tr>
<td>Monotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>FUNCTION, 2016</td>
<td>RCT</td>
<td>N=1,162</td>
<td>2 yrs*</td>
<td>TCZ (4 mg/kg monthly) + MTX (20 mg/wk) vs. TCZ (8 mg/kg monthly) + MTX vs. MTX</td>
<td>At 52 weeks, significantly greater improvement in mean HAD-DI scores from baseline in TCZ 8 mg + MTX group than in MTX group (p=0.0024)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Medium</td>
<td></td>
<td>At 24 weeks and at 52 weeks: Significantly greater change in SF-36 PCS scores in the TCZ 8 mg/kg + MTX group than in the MTX group (p=0.0014 and p=0.0066 for both time points) No differences in SF-36 PCS scores between the TCZ 4 mg/kg + MTX group and the MTX group or between TCZ and MTX group No differences in SF-36 MCS scores</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IMAGE, 2011, 2012</td>
<td>RCT</td>
<td>N=755</td>
<td>2 yrs</td>
<td>RIT (1 g days 1 and 15) + MTX (7.5-30 mg/wk) vs. RIT (500 mg days 1 and 15) + MTX vs. MTX</td>
<td>At week 52: Significantly greater improvement in physical function (measured by HAQ-DI decrease &gt;0.22) in the RIT 1 g days 1 and 15 + MTX and the RIT 500 mg days 1 and 15 + MTX groups than in the MTX group (HAQ response: 88% and 87% vs. 77%, p&lt;0.05). This difference remained for the RIT 1 g + MTX vs. the MTX group at 2 yrs (p&lt;0.05). Significantly greater improvement in the SF-36 PCS for both the RIT + MTX groups than in the MTX group (mean changes: 10.76 and 10.07 vs. 7.24, p=&lt;0.0001) Nonsignificantly greater changes in SF-36 MCS scores for both the RIT + MTX groups than in the MTX group (mean changes: 6.66 and 6.18 vs. 4.84) Significantly greater improvement in patient-reported pain (VAS, mean change) and in patient-reported fatigue (FACIT-F) in the RIT +MTX groups than in the MTX group.</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Rating</td>
<td>Study Design</td>
<td>N Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>---------------------</td>
<td>--------------</td>
<td>------------</td>
<td>------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>U-Act-Early, 2016&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>N=317</td>
<td>2 yrs</td>
<td>TCZ (8 mg/kg IV monthly) + MTX 10-30 mg/wk) vs. TCZ vs. MTX</td>
</tr>
<tr>
<td>TNF vs. Non-TNF</td>
<td>ORBIT, 2016&lt;sup&gt;1&lt;/sup&gt;</td>
<td>High</td>
<td>RCT</td>
<td>N=329</td>
<td>1 yr</td>
<td>RIT (1g days 1 and 15 and after day 26 if persistent disease activity) vs. ADA (40 mg biwkly) or ETN 50 mg/wk)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>-------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>BeSt, 2005, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2016</td>
<td>RCT</td>
<td>N=508</td>
<td>12 months (10 yrs)</td>
<td>DAS-driven treatment; G1: sequential mono-therapy starting with MTX (15 mg/week) vs. G2: stepped-up combination therapy (MTX, then SSZ, then HCQ, then PRED) vs. G3: combination with tapered high-dose PRED (60 mg/d to 7.5 mg/day) vs. G4: combination (MTX 25-30 mg/week) with IFX (3 mg/kg every 8 weeks, per DAS, could be titrated to 10 mg/kg)</td>
<td>At 3, 6, 9, and 12 months, significantly greater improvement in functional capacity in G1 and G2 vs. G3 and G4 (HAQ score improvement from baseline, p=0.05, p&lt;0.05, p&lt;0.05, and p&lt;0.05 at each time point, respectively) At 3 and 6 months, significantly greater improvement in SF-36 PCS in G1 and G2 than in G3 and G4 (p&lt;0.001); no difference in SF-36 MCS At 2 yrs, no significant differences among groups in functional capacity At 5- and 10-yr follow-up: no significant differences between groups</td>
</tr>
<tr>
<td>Medium (10-yr outcomes)</td>
<td>TEAR, 2012, 2013</td>
<td>RCT</td>
<td>N=755</td>
<td>2 yrs</td>
<td>Immediate MTX (20 mg/wk) plus ETN (50 mg/wk) vs. Immediate MTX plus SSZ (1-2 g/day) plus HCQ (400 mg/day) vs. Step up MTX to combo (MTX plus ETN) vs. Step up MTX to combo (MTX plus SSZ plus HCQ)</td>
<td>At 48 and 102 weeks, no difference in functional capacity among groups</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results (Patient-Reported Outcomes, Functioning, Quality of Life)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>GUEPARD, 2009&lt;sup&gt;12&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=65</td>
<td>1 yr</td>
<td>1: ADA 40 mg every 2 wks + MTX (max 20 mg/wk); treatment adjusted every 3 months to achieve DAS28 &lt;3.2</td>
<td>At 1 yr, no difference between groups in functional capacity, SF-36 PCS or MCS scores, pain, fatigue, or patient global assessment</td>
</tr>
<tr>
<td></td>
<td>Medium (12-wk outcomes)</td>
<td></td>
<td></td>
<td></td>
<td>2: MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High (52-wk outcomes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>OPERA, 2013&lt;sup&gt;100&lt;/sup&gt;</td>
<td>RCT</td>
<td>N=180</td>
<td>2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/wk) vs. MTX (also used intra-articular triamcinolone therapy in both groups)</td>
<td>At 1 yr, significantly greater improvement in functionality in ADA + MTX group than in MTX group (HAQ median change: -0.88 vs. -0.63, p=0.012)</td>
</tr>
<tr>
<td></td>
<td>2014&lt;sup&gt;36&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2015&lt;sup&gt;104&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2016&lt;sup&gt;162&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2017&lt;sup&gt;163&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Although the FUNCTION trial lasted a total of 2 yrs, the latest time point at which KQ 2-eligible outcomes were reported was 1 yr.

<sup>b</sup> This study evaluates comparisons in both the High-Dose Corticosteroid and TNF Biologic vs. csDMARD monotherapy categories.

<sup>c</sup> This study evaluates comparisons in both the csDMARD vs. TNF Biologic and TNF Biologic vs. csDMARD monotherapy categories.

<sup>d</sup> These studies evaluate comparisons in both the csDMARD vs. Non-TNF Biologic and Non-TNF Biologic vs. csDMARD monotherapy categories.

ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; BRAF-MDQ = Bristol Rheumatoid Arthritis Fatigue – Multidimensional Questionnaire; CI = confidence interval; csDMARD = conventional synthetic DMARD; CZP = certolizumab pegol; DAS = Disease Activity Score (based on 44 joints); DAS28 = Disease Activity Score 28; DMARD = disease-modifying antirheumatic drug; EQ-5D = EuroQol standardized instrument; ETN = etanercept; g = gram; G = group; HAQ = Health Assessment Questionnaire; HAQ-DI = Health Assessment Questionnaire-Disability Index; HCQ = hydroxychloroquine; HRQOL = health related quality of life; IFX = infliximab; IM = intramuscular; kg = kilogram; max = maximum; LEF = leflunomide; mg = milligrams; MCS = mental component score; methyl-PNL = methylprednisolone; MTX = methotrexate; N = number (of patients); NR = not reported; NS = not significant; OR = odds ratio; PCS = physical component score; PNL = prednisolone; PRED = prednisone; RCT = randomized controlled trial; RIT = rituximab; SF-12 = 12-Item Short Form Survey; SF-36 MCS = Short Form 36 Health Survey Mental Component Score; SF-36 PCS = Short Form 36 Health Survey Physical Component Score; SHS = Sharp/van der Heijde Score; SSZ = sulfasalazine; TCZ = tocilizumab; TNF = tumor necrosis factor; TOF = tofacitinib; TSU = tight step-up; VAS = visual analogue scale; vs. = versus; wk(s) = week(s); WPS-RA = Work Productivity Survey - Rheumatoid Arthritis; yr(s) = year(s).
Detailed Synthesis

Corticosteroids

Corticosteroids Versus csDMARDs

Evidence was insufficient to determine whether patients treated with corticosteroids plus csDMARDs versus csDMARD monotherapy differed on functional capacity or HRQOL.

Five RCTs (n=1,329 eligible) compared a combination of a corticosteroid plus a csDMARD with csDMARD only and were eligible for this Key Question; four examined functional capacity or quality-of-life outcomes (or both), and one examined patient-reported symptoms only (Table 9). Two studies added prednisolone (PNL) to either MTX or SSZ, two studies examined adding prednisone (PRED) to MTX, two studies added PRED to SSZ, and one study added PRED to leflunomide (LEF).

The duration and dose of PRED varied among studies. Doses ranged from 7.5 mg per week to taper schedules starting at 60 mg per week. The duration and dosing of PNL also varied, with a dose of 7.5 mg per day in one study and a taper schedule starting at 60 mg per day in another. Overall, improvements in functional capacity were mixed. Three studies demonstrated significant improvements and one showed no difference.

In the CAMERA-II trial, functional capacity as measured by HAQ mean difference improved significantly more at 2 years in the PNL plus MTX group than in the MTX monotherapy group (HAQ mean difference, -0.18; 95% CI, -0.34 to -0.02) (p=0.027). It should be noted that the difference of at least 0.20 is considered to represent a clinically significant change (Appendix F). Similar statistically significant differences were found at 3, 6, 12, and 18 months. In the BARFOT #2 trial, physical function as measured by mean decrease in HAQ improved significantly more from baseline in the PNL plus csDMARD group than in the csDMARD monotherapy group at 3, 6, 12, and 18 months and 2 years (p=0.003); the difference was still present in the followup at 4 years (p=0.034). In the CARDERA trial, at 2 years, functional capacity did not differ between the PNL plus MTX group and MTX monotherapy group (HAQ mean change, -0.28 vs. -0.27, p=NR, respectively). In the CareRA trial, functional capacity did not differ among the groups at 16 weeks and 54 weeks as measured by clinically meaningful change in HAQ. In the CareRA trial, functional capacity did not differ significantly among the groups at 16 weeks and 54 weeks as measured by clinically meaningful change in HAQ.

One RCT evaluated HRQOL outcomes. The investigators found no significant differences between PNL plus MTX and MTX monotherapy in either the physical or the mental subscale of the 36-Item Short-Form Health Survey (SF-36) or the EuroQoL standardized instrument (EQ-5D) (p=0.22).

One RCT evaluated patient-reported symptoms and found significantly greater improvement in pain as measured with a Visual Analogue Scale (VAS) in the PRED plus MTX group compared with the MTX monotherapy group at 4 months (p=0.01) and 12 months (p=0.04).

High-Dose Corticosteroids

Two RCTs evaluated the efficacy of high-dose corticosteroids in MTX-naïve populations. In the IDEA trial (N=112), a single high dose of methyl-PNL (250 mg IV) plus MTX was compared with IFX plus MTX over 26 weeks with a 50-week open-label extension. Groups did
not differ in functional capacity at 26 and 78 weeks, as measured by mean change in HAQ-Disability Index [DI] (at 78 weeks, -0.85 vs. -0.79; p=0.826). The second study (N=44)\(^\text{18}\) compared IFX plus MTX versus high-dose methyl-PNL (1 g IV at weeks 0, 2, and 6 and then every 8 weeks for 46 weeks) plus MTX versus MTX monotherapy. At 52 weeks, this study found significantly greater HAQ improvements over time in the methyl-PNL plus MTX group than in the MTX group (p=0.001).

**csDMARDs**

**csDMARDs Versus csDMARDs**

**csDMARD Monotherapy Versus csDMARD Monotherapy**

One RCT (N=245) compared MTX plus PNL with SSZ plus PNL. Functional capacity did not differ significantly at 2 years between groups (HAQ mean change from baseline, -0.35 vs. -0.38; p=0.752).\(^\text{27}\)

One observational study compared SSZ (2 g/d) with MTX (10-15 mg/wk) monotherapy. At 6 months, functional capacity improved significantly in the MTX group compared with the SSZ group (modified HAQ mean change from baseline, -0.26 vs. -0.13; p=0.002).\(^\text{28}\) However, this difference was not significant after adjusting for propensity score quintile and physician global VAS. HRQOL outcomes did not differ between groups as measured by mean change from baseline values on the SF-36 physical and mental component subscales. There was no significant difference in patient-reported pain or fatigue as measured by VAS mean change from baseline between groups. Of note, both the RCT and observational study used MTX dosing that is lower (5-15 mg weekly) than typically recommended as efficacious (20-25 mg weekly).

**csDMARD Combination Therapy Versus csDMARD Monotherapy**

Six RCTs (N=1,347) compared combination csDMARD therapy with csDMARD monotherapy. Four trials examined the combination SSZ plus MTX versus csDMARD monotherapy (MTX or SSZ).\(^\text{21, 23-25}\) Two other trials examined the combination of MTX plus SSZ plus HCQ against csDMARD monotherapy with different PRED doses.\(^\text{4, 22}\) Trial durations ranged from 1 to 5 years. Doses of MTX were variable, ranging from 7.5 mg weekly to 25 mg weekly.

All six trials found no significant differences in functional capacity between the combination csDMARD group and the csDMARD monotherapy at 1 to 5 years.\(^\text{4, 21-25}\) One trial found significant improvement in functional capacity in the combination csDMARD group at 28 weeks, measured as a mean change in HAQ (-1.1 vs. -0.6, p<0.0001), but this difference was not sustained at either 52 weeks or 5 years.\(^\text{24}\) This same trial found greater improvement in patient-reported pain (VAS, mean change -34 vs. -20, p<0.002) in the combination csDMARD group compared with the csDMARD monotherapy group at 28 weeks but no difference between groups at 56 weeks. One trial\(^\text{148}\) found no difference in quality of life over time, measured with the EQ-5D, between the csDMARD combination group and the csDMARD monotherapy group. In the FIN-RACo study,\(^\text{22}\) patients treated with MTX plus SSZ plus HCQ plus PNL had significantly less work disability at 2 years than patients receiving csDMARD monotherapy (MTX or SSZ) (median work disability per patient-observation years, in days: 12.4 vs. 32.2; p=0.008). In the tREACH trial, patients treated with MTX plus SSZ plus HCQ plus glucocorticoids had less unemployment than patients receiving MTX plus glucocorticoids at 12 months (p=0.015).
csDMARDs Versus Biologics

**TNF Biologic: MTX Plus TNF Biologic Versus Monotherapy With Either MTX or TNF Biologic**

The PREMIER study (N=799) examined the combination of ADA (40 mg biwkly) plus MTX (20 mg/wk) compared with either ADA alone or MTX alone in patients with early aggressive RA. At 1 year, the ADA plus MTX group achieved significantly greater improvement in functional capacity than the ADA group (HAQ-DI mean change: -1.1 and -0.8, respectively; p=0.0002).

At 2 years, several outcomes appeared to favor the combination groups. The ADA plus MTX group had more improvement in functional capacity than the MTX group (HAQ-DI mean change, -1.0 vs. -0.9; p≤0.05). Additionally, significantly more patients in the ADA plus MTX group had a HAQ-DI score of 0 than did those in either monotherapy group (33% vs. 19% vs. 19%; p<0.001). The ADA plus MTX group had a greater improvement in quality-of-life outcomes than the MTX group based on the physical subscale of the SF-36 (PCS) but not the mental subscale (MCS); the ADA-only group had statistically higher improvements than the MTX-only group based on the SF-36 MCS (p=0.0148). The ADA plus MTX group had lower patient-reported pain (mean pain VAS) than the ADA-only group (9.6 vs. 19.6, p<0.0001). There was no difference in patient-reported pain between the ADA-only group and the MTX-only group. Finally, compared with patients in the MTX-only group, patients in the ADA plus MTX group had more gained employment (27.4% vs. 22.7%) and fewer missed work days (mean 17.4 for 130 employed vs. 36.9 for 110 employed).

**Non-TNF Biologic: MTX Plus Non-TNF Biologic Versus Monotherapy With Either MTX or Non-TNF Biologic**

One trial, the multinational AVERT trial (n=351), compared the combination of ABA (125 mg/week subcutaneous) plus MTX (7.5 mg/week) with ABA monotherapy. This double-blind RCT compared treatments over 1 year; at year 2, patients with a DAS28-CRP <3.2 were tapered off treatment. If patients had an RA flare by month 15, they were given ABA plus MTX. The percentage of patients who had HAQ-DI response in the ABA plus MTX group was higher than the percentages in the ABA group at 12 months (65.5% vs. 52.6%) and 18 months (21.8% vs. 16.4%), but these differences were not statistically significant.

Two RCTs compared the combination of TCZ plus MTX with TCZ alone or MTX alone. Both trials demonstrated greater functional capacity in the combination TCZ (8 mg/kg) and MTX group than in the TCZ-alone or MTX-alone groups.

In the FUNCTION trial (N=1,162), the TCZ (8 mg/kg) plus MTX group achieved a statistically greater improvement in functional capacity than the MTX group (mean change from baseline HAQ-DI -0.81 vs. -0.64 p=0.0024) at 52 weeks. A significantly greater improvement in SF-36 PCS was seen in the TCZ (8 mg/kg) plus MTX group than in the MTX group at 24 weeks (p=0.0014) and at 52 weeks (p=0.0066). By contrast, functional capacity or HRQOL did not differ between the TCZ (4 mg/kg) plus MTX and MTX groups or between TCZ monotherapy and MTX monotherapy groups at either 24 or 52 weeks.

The U-Act-Early trial (N=317) used the Dutch HAQ to assess physical function. Significantly greater improvement in functional capacity was demonstrated at 24 weeks in the combination TCZ plus MTX group than in the TCZ-alone or the MTX-alone group at 24 weeks (p=0.0275). This difference was not found at 52 or 104 weeks. Additionally, there was significantly greater improvement in mean SF-36 PCS scores over time in the TCZ plus MTX
group and TCZ-alone group than in the MTX-alone group (p=0.044 and p=0.012, respectively). No significant differences were found in SF-36 MCS scores over time between groups. There was also significantly greater improvement in mean EQ-5D scores over time in the TCZ plus MTX group than in the MTX-alone group (p=0.018). There was no significant difference between the TCZ-alone and MTX-alone groups.135

**csDMARDs Versus tsDMARDs: MTX Plus tsDMARD Versus Either MTX or tsDMARD**

One RCT examined (N=108) the combination of TOF (20 mg/day, higher than the dose typically used) plus MTX (10-20 mg/week) against TOF alone or MTX alone in patients with early active RA.29 It found no significant difference across these groups in functional capacity improvement, as measured by HAQ-DI improvement from baseline >0.22, at 3, 6, or 12 months.29

**Biologics**

**TNF Biologics**

**TNF Biologic Versus csDMARD Monotherapy**

Thirteen RCTs examined whether adding a TNF biologic improved outcomes in csDMARD users. The TNF biologics included were ADA, CZP, ETN, and IFX. No eligible trial or study was found for GOL. All involved a csDMARD (typically MTX) as the comparison group. The time frames of these trials differed considerably. Most of our 13 trials suggested greater improvement in functional capacity with a combination TNF biologic and csDMARD than with csDMARD monotherapy.12, 13, 15-18, 34, 35, 37, 41, 103, 114-119, 150-152 This finding applied to the following TNF biologics: ADA (difference of HAQ change -0.1 to -0.3 over 24 weeks to 2 years) (moderate SOE), CZP (difference of HAQ change not consistently reported, but in favor of combination therapy, over 30 weeks to 1 year) (low SOE), and IFX (difference of HAQ change not consistently reported, but in favor of combination therapy, over 30 weeks to 1 year) (low SOE). Evidence was insufficient to determine the impact on HRQOL of adding TNF biologics to MTX therapy. The results of the trials reporting HRQOL outcomes were mixed. Several trials demonstrated improvement in SF-36 PCS scores;12, 17, 34, 36 none showed improvement in other measures.

One trial comparing ETN monotherapy with MTX monotherapy showed no significant difference in mean HAQ scores at 12 months but greater improvement in functional capacity at 24 months in the ETN monotherapy group (open-label extension).14

**Adalimumab.** Five RCTs compared ADA (40 mg biweekly) plus MTX (ranging from 8 to 20 mg/week) with MTX monotherapy.13, 15, 16, 34-37, 103, 114-119, 150-152, 160-163 The HIT HARD trial demonstrated clinically significantly greater functional capacity in the ADA and MTX group than in the MTX group at 24 weeks (mean HAQ-DI, 0.49 vs. 0.72; p=0.0014).34 At 24 weeks, scores on the SF-36 PCS were significantly higher for ADA plus MTX patients than MTX-only patients (44 vs. 39.8, p=0.0002) but patients in these two groups did not differ on the SF-36 MCS. At 48 weeks, the trial detected no differences in functional capacity and HRQOL.

In the HOPEFUL 1 trial,35 the ADA plus MTX group experienced a clinically significant larger improvement in physical function than the MTX group (decrease from baseline mean HAQ-DI score, 0.6 vs. 0.4; p<0.001); in addition, significantly more patients in the ADA plus
MTX than in the MTX group achieved normal functionality (HAQ-DI score <0.5, 60.0% vs. 36.8%; p=0.001) at 26 weeks.

The OPTIMA trial was a phase 4 multinational RCT comparing ADA plus MTX with MTX in early RA. At 26 weeks, the study demonstrated clinically significant greater functional improvements in the ADA plus MTX group than in the MTX group (HAQ-DI mean score, 0.7 vs. 0.9; p<0.001); in addition, a significantly greater proportion of ADA plus MTX patients than MTX-only patients demonstrated normal function (40.0% vs. 28.0%, respectively; p<0.001). In post hoc analysis, the ADA plus MTX group had significant improvement in work-related outcomes at 26 weeks compared with the outcomes in the MTX group (patients receiving ADA plus MTX showed significant changes in percentage points from baseline compared with patients receiving MTX in activity impairment, presenteeism, and overall work impairment [32.0% vs. 23.7%, 24.6% vs. 17.1%, 27.3% vs. 18.3%, respectively]). In patients who had achieved low disease activity at 26 weeks, the two therapy groups did not differ in physical functional score at 78 weeks.

The PREMIER study (N=799), also described previously in the csDMARDs versus Biologics section, examined the combination of ADA plus MTX compared with MTX alone in patients with early aggressive RA. At 1 year, the ADA plus MTX group achieved clinically significant greater improvement in functional capacity than the MTX group (p=0.0003) (HAQ-DI mean change: -1.1 vs. -0.8).

In the PROWD study, the primary outcome was to evaluate work disability in each group. At week 16, fewer patients in the ADA plus MTX group than in the MTX group had job loss, (16% vs. 27.3%, p=0.092). At 56 weeks, job loss was significantly lower with ADA plus MTX compared with MTX (18.6% vs. 39.7%, p<0.005). At 56 weeks, the ADA plus MTX patients had significantly greater improvement in function from baseline than the MTX patients (change in mean HAQ, -0.7 vs. -0.4; p=0.005).

Certolizumab. Two RCTs examined the combination of CZP (either 400 mg biweekly for 4 weeks or 200 mg biweekly for 4 weeks, then 200 mg biweekly) plus MTX with MTX only. The C-OPERA trial randomized 316 patients with early RA with poor prognostic factors (high anti-CCP antibody, positive RF or bony erosions). The CZP plus MTX group experienced a rapid and statistically significant (p<0.05) improvement in HAQ-DI response rate compared with the MTX group at all time points from 4 weeks to 52 weeks. At 104 weeks, HAQ remission rates were higher in the CZP plus MTX group compared with the MTX group but did not meet statistical significance (73% vs. 63.7%, p=0.09). The C-EARLY trial compared CZP plus MTX with MTX alone in 879 patients with early RA and poor prognostic factors (positive anti-CCP antibody or positive RF) and found a similarly significant greater improvement in functional capacity in the CZP plus MTX group than in the MTX group at 1 year (mean change in HAQ-DI from baseline, -1.00 vs. -0.82, p<0.001). The CZP plus MTX group also had greater improvement in household and work productivity than the MTX group at 52 weeks based on a work productivity scale for RA (WPS-RA). CZP plus MTX patients reported greater improvements versus MTX in household productivity (household work days missed per month baseline vs. week 52: MTX=10.4 vs. 3.0, CZP + MTX=8.8 vs. 1.9; household work days with productivity reduced by ≥50%/month: MTX=10.6 vs. 3.0, CZP + MTX=9.4 vs. 2.1; level of arthritis interference with household work productivity/month: MTX=6.4 vs. 2.5, CZP + MTX=6.0 vs. 1.9). Employed CZP plus MTX patients reported reductions in absenteeism and increases in presenteeism versus MTX (work days missed per month, baseline vs. week 52: MTX=4.0 vs. 0.9, CZP + MTX=4.4 vs. 0.6; days with work productivity reduced per month:...
MTX=8.8 vs. 1.8, CZP + MTX=6.4 vs. 1.0; level of arthritis interference with work productivity/month: MTX=5.8 vs. 1.9, CZP + MTX=5.5 vs. 1.4).

*Etanercept.* Three RCTs compared ETN (25 mg twice weekly or 50 mg weekly) with MTX. The COMET trial compared ETN plus MTX with MTX alone. It found a clinically significant greater improvement in functional capacity in the ETN plus MTX group than in the MTX group at 52 weeks (HAQ mean change: -1.02 vs. -0.72, p<0.0001). Significantly more patients in the ETN plus MTX group than in the MTX group achieved normal function (HAQ-DI<0.5) (55% vs. 39%, p=0.0004) at 52 weeks. They also had significantly higher SF-36 PCS scores (13.7 vs. 10.7, p=0.003), but did not differ from the MTX group in the SF-36 MCS scores. In post hoc analysis, improvement in work-related outcomes was apparent; significantly fewer patients had to stop working (8.6% vs. 24%, p=0.004) and fewer had problems with absenteeism (mean missed workdays: 14.2 vs. 31.9).

In the Enbrel Early RA study, ETN 25 mg twice weekly was compared with MTX over 12 months. Physical function did not differ between groups (~55% in each arm had at least a 0.5-unit improvement in HAQ) at 12 months. In the open-label extension from 12 to 24 months, significantly more patients in the ETN group than in the MTX group achieved improvement in function (HAQ improvement >0.5 units: 37% vs. 55%, p<0.001).

A smaller trial (n=26) compared ETN 25 mg twice weekly with MTX over 24 weeks and found greater improvement in function in the ETN group than in the MTX group at 12 weeks (HAQ mean change from baseline, 0.9 vs. 0.6; p=NR) but no further improvement seen in either group from 12 to 24 weeks (p=0.38).

*Infliximab.* Three trials compared the combination of IFX plus MTX with MTX monotherapy.

The ASPIRE trial (n=1,049) was a 54-week trial comparing IFX (3 mg/kg or 6 mg/kg) plus MTX with MTX monotherapy. More patients in the IFX 3 mg/kg and 6 mg/kg + MTX groups than the MTX group had clinically significant improvements in HAQ scores from baseline to 54 weeks (percentage of patients with HAQ increase ≥0.22 units from baseline: 76%, 75.5%, 65.2%; p<0.004). The average improvement in physical function from 30 to 54 weeks was significantly greater in the IFX 6 mg/kg plus MTX and IFX 3 mg/kg plus MTX groups than in the MTX monotherapy group (mean decrease in HAQ scores from baseline: 0.88, 0.80, vs. 0.68, p<0.001). At 54 weeks, HRQOL ratings (SF-36 PCS score) were significantly higher in both IFX plus MTX groups than in the MTX group (11.7, 13.2, vs. 10.1; p=0.003). Additionally, this study assessed work disability by patient-reported working capacity, or employability, at baseline and 54 weeks. For this analysis, IFX 3 mg/kg and 6 mg/kg groups were combined. Employability improved significantly in the IFX plus MTX group compared with those outcomes in the MTX group (employability odds ratio [OR] [95% CI]: 2.4 [2.2 to 2.6]; p<0.001) and significantly fewer patients were unemployable (8% vs. 14%, p=0.05). By contrast, it found no significant differences in the change in employment rates between the IFX plus MTX group and the MTX group (0.5% vs. 1.3%; p>0.05). Of note, work disability was a secondary outcome measure in the study.

One small trial (n=20) also found a significant functional benefit (by HAQ) at 54 weeks favoring IFX (3 mg/kg at standard intervals) plus MTX over MTX (p<0.05). In the 8-year followup, physical function outcomes did not differ between groups (HAQ median [IQR]: 1.0 [0.1-1.8] vs. 1.5 [1.2-2.1]; p=0.12).

Another small trial (n=44), also described previously in the High-Dose Corticosteroids section, compared IFX 3 mg/kg plus MTX with MTX alone over 1 year. Although the IFX plus
MTX group experienced a significant improvement in functional capacity (by HAQ) over time, its change in functional capacity did not differ significantly compared with the MTX group (p=NR).

**TNF Biologic Versus csDMARD Combination Therapy**

The TNF biologic IFX plus the FIN-RACo regimen (a combination of csDMARDs - MTX, HCQ, and SSZ – plus PRED) versus the FIN-RACo regimen alone did not differ significantly in their impact on functional capacity (low SOE). Three RCTs examined the impact of TNF biologics compared with csDMARD combination therapy. One trial evaluated ADA; two trials evaluated IFX. Two trials reported functional capacity outcomes; they reported no significant difference in physical function between groups at all time points ranging from 4 months to 5 years. Two studies examined quality-of-life outcomes and found no significant differences between groups. One study examined patient-reported pain and found significantly lower patient-reported pain in the ADA plus MTX group compared with the combination csDMARD group at 1 year (mean pain VAS, 28 vs. 38, p=0.02) and no significant difference at 8 months. Evidence was insufficient to determine the impact of the TNF biologic ADA or IFX plus MTX versus csDMARD triple therapy on functional capacity.

**Non-TNF Biologics**

**Non-TNF Biologic Plus MTX Versus MTX Monotherapy**

*Abatacept.* Two RCTs evaluated the combination of ABA plus MTX in comparison with MTX alone. The AGREE trial compared the ABA (10 mg/kg IV) plus MTX (7.5 mg/week) group with the MTX group over 2 years. We rated this trial as high ROB because overall discontinuation rates were high (up to 42 percent). The first year was a double-blind trial; in year 2, patients in the ABA plus MTX group continued treatment and patients in the MTX-only group were started on ABA. At 1 year, the ABA plus MTX patients had clinically significant greater functional benefit than patients in the MTX group (HAQ-DI % change of >0.3 units from baseline: 71.9% vs. 62.1%, p=0.024). Significant improvements in quality-of-life outcomes occurred in the ABA plus MTX group compared with outcomes in the MTX group; these were assessed by mean changes from baseline in the SF-36 MCS (8.15 vs. 6.34, p=0.046) and the SF-36 PCS (11.68 vs. 9.18, p=0.005).

The multinational AVERT trial (n=351), previously described in the csDMARDs versus non-TNF biologics section, also compared the combination of ABA (125 mg/week subcutaneous) plus MTX (7.5 mg/week) with ABA monotherapy or MTX monotherapy. This double-blind RCT compared treatments over 1 year; at year 2, patients with a DAS28-CRP <3.2 were tapered off treatment. If patients had an RA flare by month 15, they were given ABA plus MTX. The percentage of patients in the ABA plus MTX group was higher than the percentages in the MTX group who had HAQ-DI response at 12 months (65.5% vs. 44%) and 18 months (21.8% vs. 10.3%), but these differences were not statistically significant.

*Rituximab.* One RCT, the IMAGE trial (n=755), compared RIT (1 g on days 1 and 15) plus MTX (7.5 mg-20 mg/week) combination therapy, RIT (500 mg on days 1 and 15) plus MTX (7.5 mg to 20 mg/week) combination therapy, and MTX monotherapy over 2 years. At week 52, functional capacity (measured by HAQ-DI decrease >0.22) improved more in the RIT 1 g plus MTX and the RIT 500 mg plus MTX groups than in the MTX-only group (HAQ response, 88% and 87% vs. 77%; p<0.05). This difference remained for the RIT 1 g plus MTX group versus the MTX-only group at 104 weeks (p<0.05). The improvement in SF-36 PCS
scores in both RIT plus MTX groups was significantly greater than in the MTX monotherapy group (mean changes in PCS scores, 10.76 and 10.07 vs. 7.24; p <0.0001). The mean changes in SF-36 MCS were not significantly different (6.66 and 6.18 vs. 4.85). There was also significantly greater improvement in patient-reported pain in the RIT plus MTX groups than in the MTX monotherapy group (VAS, mean change, p<0.0001) and in patient-reported fatigue (FACIT-F, mean change, p<0.05) at 52 weeks.

*Tocilizumab.* Two RCTs, also described previously in the csDMARDs versus Biologics section, compared the combination of TCZ plus MTX with MTX alone.32, 33 Both trials demonstrated greater functional capacity in the combination TCZ (8 mg/kg) and MTX group than in the MTX-alone group.

In the FUNCTION trial (N=1,162),32 the TCZ (8 mg/kg) plus MTX group achieved a statistically greater improvement in functional capacity than the MTX group (mean change from baseline HAQ-DI -0.81 vs. -0.64, p=0.0024) at 52 weeks. A significantly greater improvement in SF-36 PCS was seen in the TCZ (8 mg/kg) plus MTX group than in the MTX group at 24 weeks (p=0.0014) and at 52 weeks (p=0.0066). By contrast, functional capacity or HRQOL did not differ between the TCZ (4 mg/kg) plus MTX and MTX groups at either 24 or 52 weeks.

The U-Act-Early trial (N=317) used the Dutch HAQ to assess physical function.33 Significantly greater improvement in functional capacity was demonstrated at 24 weeks in the combination TCZ plus MTX group than in the MTX-alone group at 24 weeks (p=0.0275). This difference was not found at 52 or 104 weeks. Additionally, there was significantly greater improvement in mean SF-36 PCS scores over time in the TCZ plus MTX group than in the MTX-alone group (p=0.044). No significant differences were found in SF-36 MCS scores over time between groups. This trial also found significantly greater improvement in mean EQ-5D scores over time in the TCZ plus MTX group than in the MTX-alone group (p=0.018). There was no significant difference between the TCZ-alone and MTX-alone groups.135

**Biologic Head to Head: TNF Versus Non-TNF**

Evidence was insufficient to determine any differences between one biologic and another biologic for either the functional capacity or the HRQOL outcomes. One RCT compared TNF biologics with non-TNF biologics. The ORBIT trial, an open-label noninferiority RCT (n=329) over 1 year, compared the non-TNF RIT (1 g days 1 and 15) with TNF treatment (either ADA (40 mg biweekly) or ETN (50 mg/week)).8 Patients had had a prior inadequate response to at least two csDMARDs. Patients in the RIT group had a statistically greater improvement in physical function (mean HAQ change from baseline) than in the TNF group at 6 months (-0.44 vs. -0.31; p=0.0391) and 12 months (-0.49 vs. -0.38; p=0.0391). The EQ-5D and anxiety and depression measures did not differ at 6 months and 12 months.

**Combinations and Therapy Strategies**

Combination strategies using multiple csDMARDs or csDMARD plus TNF biologics compared with sequential or step-up therapies did not differ significantly in terms of functional capacity (low SOE). Evidence is insufficient to determine the impact of these strategies on HRQOL. Two RCTs20, 83, 85, 159 evaluated combination strategies using corticosteroids plus oral DMARDs or TNF biologics. The results of these studies demonstrated that using combination therapy produced significantly more rapid improvement in functional capacity (difference in mean change in HAQ at 28 weeks, -0.5; p<0.0001) and less work disability (median, 12.4 days per patient-observation year vs. 32.3 days; p<0.008) than oral DMARD monotherapy.
The BeSt RCT examined four different treatment strategies over 12 months. Patients treated with initial combination csDMARD therapy plus PRED (group 3) or initial combination therapy plus IFX (group 4) had more rapid improvement in functional ability than those treated with sequential csDMARD therapy (group 1) or with step-up combination therapy (group 2). Statistically significant improvements were reported for 3, 6, 9, and 12 months. By 2 years, all groups maintained their improvements but the groups themselves did not differ significantly. Improvements were also maintained at 4-, 5-, and 10-year followup. Patients in groups 3 and 4 also had more rapid improvement in physical HRQOL, with greater improvements at 3 months and 6 months for groups 3 and 4 than for groups 1 and 2 on the SF-36 PCS (p<0.001). By years 1 and 2, all groups had similar improvement in SF-36 PCS. Mental HRQOL measured by the SF-36 MCS did not differ across groups.

The TEAR study found no significant difference in functional ability at 48 or 102 weeks. The comparisons were four groups: immediate combination TNF biologic and csDMARD group (group 1); immediate combination csDMARD group (group 2); step-up from MTX to TNF biologic plus MTX (group 3); and step-up from MTX to combination csDMARD group (group 4).

The GUEPARD study compared the initial strategy of ADA (40 mg every 2 weeks) plus MTX (up to 20 mg/wk) with MTX monotherapy for 3 months. In patients who did not respond to an initial strategy at 3 months, the investigators examined whether a disease activity–driven treatment strategy with TNF biologics was equally effective in both groups. At 1 year, there was no difference between groups in functional capacity, SF-36 PCS, or SF-36 MCS scores. There was no difference between groups in patient-reported pain or fatigue at 12 weeks or 1 year. Of note, this study was rated high ROB after 3 months because of the risk of contamination bias based on modifications in treatment dosing and regimens when low disease activity was achieved.

The OPERA trial of 180 Danish early RA patients compared ADA (40 mg every 2 weeks) plus MTX (7.5 mg-20 mg) with MTX alone. At 3 months, SSZ or HCQ could be added if disease activity persisted. There was a clinically significant greater improvement in functional capacity at 1 year in patients treated with initial combination therapy (ADA plus MTX) than in monotherapy (MTX) patients (HAQ median change: -0.88 vs. -0.63; p=0.012). The improvement in the SF-12 PCS was also greater for the combination than the monotherapy patients (13.2 vs. 10.6; p=0.0150), and the combination group also reported significantly less pain (median VAS score, p=0.007), but there were no differences in change in the SF-12 MCS, the EQ-5D, or fatigue. At 2 years, the groups did not differ in physical function, quality of life, pain, or fatigue.

KQ 3: Comparative Harms of Drug Therapies for Patients With Early RA in Relation to Harms, Tolerability, Patient Adherence, or Adverse Effects

For this KQ, we use the FDA definition for serious adverse events. These include death, life-threatening experience, hospitalization or prolongation of hospitalization, significant incapacity or inability to conduct normal life functions, congenital anomaly, medical event requiring medical or surgical intervention to prevent one of the prior outcomes. Specific adverse events include 11 most commonly occurring across all our eligible drugs according to their FDA-approved labels. This set of adverse events includes rash, upper respiratory tract infection,
nausea, pruritus, headache, diarrhea, dizziness, abdominal pain, bronchitis, leukopenia, and injection site reactions.

**Key Points**

- Conclusions below are based on early RA studies including patients with moderate to high disease activity, and the majority were MTX naive.
- Clinical trials provided the majority of evidence that was available for this population.
- Corticosteroids and csDMARDs did not differ significantly in serious adverse events (moderate SOE) or discontinuation rates attributable to adverse events (low SOE).
- csDMARD combination therapy compared to csDMARD monotherapy did not differ significantly in serious adverse events (low SOE). Combining a csDMARD with a TNF biologic did not differ significantly in serious adverse events (moderate SOE) or discontinuations attributable to adverse events compared with TNF biologic monotherapy (moderate SOE). Similarly, combining a csDMARD with a non-TNF biologic did not lead to a significant difference in serious adverse events (moderate SOE) or discontinuations attributable to adverse events compared with non-TNF biologic monotherapy (moderate SOE).
- Serious adverse events or discontinuations attributable to adverse events did not differ significantly between the TNF biologics (ADA, CZP, ETN, IFX) in combination with MTX versus MTX monotherapy (low SOE).
- Discontinuations attributable to either adverse events or serious adverse events did not differ significantly between the non-TNF biologics (ABA, RIT, TCZ) in combination with MTX versus MTX monotherapy (low SOE for ABA and moderate SOE for RIT and TCZ).
- Harms evidence was insufficient for head-to-head comparisons of TNF and non-TNF biologics.
- Long-term studies (up to 10 years) of combination strategies using multiple csDMARDs or csDMARD plus TNF biologics ultimately showed no differences in serious adverse events between immediate combination and step-up therapies (low SOE).

**Detailed Synthesis**

Table 10 presents data on all included trials or observational studies for the four main outcomes of concern for KQ 3: overall discontinuation rates; discontinuations attributable to adverse events; serious adverse events; and occurrence of specific adverse events. All outcomes were reported in percentages.
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr</th>
<th>Study Design</th>
<th>N</th>
<th>Duration</th>
<th>Comparison (Dose)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CAMERA-II, 2012&lt;br/&gt;Medium&lt;br/&gt;RCT&lt;br/&gt;N=239&lt;br/&gt;2 yrs</td>
<td>PRED (10 mg/day) + MTX (10 mg/week) vs. MTX (10 mg/week)</td>
<td>Overall discontinuation: 28% vs. 29.8% at 2 years&lt;br/&gt;Discontinuation due to adverse events: 14% vs. 17%&lt;br/&gt;Serious adverse events: 2.0% vs. 4.0%&lt;br/&gt;Specific adverse events:&lt;br/&gt;Nausea: 19.6% vs. 36.1, p=0.006&lt;br/&gt;ALT &gt; ULN: 12.8% vs. 27.7%, p=0.016&lt;br/&gt;AST &gt; ULN: 6.8% vs. 17.6%, p=0.016&lt;br/&gt;Headache: 19.6% vs. 26%&lt;br/&gt;No difference in infections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CARDERA, 2007&lt;br/&gt;Medium&lt;br/&gt;RCT&lt;br/&gt;N=467&lt;br/&gt;2 yrs</td>
<td>PNL (60 mg/day tapered over 34 weeks) + MTX (7.5-15 mg/week) vs. MTX</td>
<td>Overall discontinuation: 47% vs. 16.2% at 2 years&lt;br/&gt;Discontinuation due to adverse events: 12.2% vs. 6.8%&lt;br/&gt;Serious adverse events: 19.0% vs. 21.0%&lt;br/&gt;Specific adverse events:&lt;br/&gt;Respiratory tract infection: 49.0% vs. 54.0%&lt;br/&gt;Nausea/vomiting: 20.0% vs. 15.0%&lt;br/&gt;Abdominal pain: 9.0% vs. 7.0%&lt;br/&gt;Headache: 10.0% vs. 6.0%&lt;br/&gt;Dizziness: 6.0% vs. 4.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>Montecucco et al., 2012&lt;br/&gt;Medium&lt;br/&gt;Open label&lt;br/&gt;RCT&lt;br/&gt;N=220&lt;br/&gt;12 months</td>
<td>PRED (12.5 mg/day for 2 weeks then taper to 6.25 mg/day) + MTX (10-25 mg/week) vs. MTX (10-25 mg/week)</td>
<td>Overall discontinuation: 8.2% vs. 10.9%&lt;br/&gt;Discontinuation due to adverse events: 5.5% vs. 9.1%, p=0.29&lt;br/&gt;Serious adverse events: NR&lt;br/&gt;Specific adverse events: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Risk of Bias Ratings</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>-----------------------------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>CareRA, 2015, 2015, 2017</td>
<td>Medium</td>
<td>Open label RCT</td>
<td>379</td>
<td>2 yrs</td>
<td>High-risk patients: 1: MTX (15 mg/week) + SSZ (2 g/day) + PRED (60 mg/day tapered to 7.5 mg/day) vs. 2: MTX + PRED (30 mg tapered to 5 mg/day) vs. 3: MTX + LEF (10 mg/day) + PRED (30 mg tapered to 5 mg/day) vs. Low-risk patients: 4: MTX 15 mg/week vs. 5: MTX + PRED (30 mg tapered to 5 mg/day)</td>
</tr>
<tr>
<td>Corticosteroids vs. csDMARDs</td>
<td>BARFOT #2, 2005, 2009, 2014</td>
<td>Medium</td>
<td>Open label RCT</td>
<td>259</td>
<td>2 yrs</td>
<td>PNL 7.5 mg/day + DMARD (SSZ 2 g/day or MTX 10 mg/week) vs. DMARD (SSZ 2 g/day or MTX 10 mg/week)</td>
</tr>
<tr>
<td>High-Dose Corticosteroids</td>
<td>Durez et al., 2007</td>
<td>a</td>
<td>RCT</td>
<td>44</td>
<td>1 yr</td>
<td>IFX (3 mg/kg at weeks 0, 2, 6 until 46 weeks) + MTX (7.5-20 mg/wk) vs. Methyl-PNL (1 g weeks 0, 2, 6 and every 8 weeks until 46 weeks) + MTX vs. MTX</td>
</tr>
<tr>
<td>csDMARD Monotherapy Versus csDMARD Monotherapy</td>
<td>BARFOT #1, 2003</td>
<td>High</td>
<td>RCT</td>
<td>245</td>
<td>2 yrs</td>
<td>PNL (7.5-15 mg/day for 1-3 months) + MTX (5-15 mg/week) vs. SSZ (2-3 g/day) + PNL (up to 10 mg/day)</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr Risk of Bias Ratings</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Monotherapy Versus csDMARD Monotherapy</td>
<td>NOR-DMARD 2012 High</td>
<td>Observational N=1,102 3 yrs</td>
<td>MTX (10 mg-15 mg/week) vs. SSZ (2 g/day)</td>
<td>Overall discontinuation: 48.1% vs. 78.9% Discontinuation due to adverse events: 15.4% vs. 36% Serious adverse events: NR Specific adverse events: Infections: 34.1% vs. 20.0%, p&lt;0.001 Nausea: 18.9% vs. 13.1%, p&lt;0.07 Abdominal pain: 4.0% vs. 8.0%, p&lt;0.03 Rash: 2.7% vs. 9.1%, p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Monotherapy Versus csDMARD Monotherapy</td>
<td>NOR-DMARD 2012 High</td>
<td>Observational N=1,102 3 yrs</td>
<td>MTX (10 mg-15 mg/week) vs. SSZ (2 g/day)</td>
<td>Overall discontinuation: 48.1% vs. 78.9% Discontinuation due to adverse events: 15.4% vs. 36% Serious adverse events: NR Specific adverse events: Infections: 34.1% vs. 20.0%, p&lt;0.001 Nausea: 18.9% vs. 13.1%, p&lt;0.07 Abdominal pain: 4.0% vs. 8.0%, p&lt;0.03 Rash: 2.7% vs. 9.1%, p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Dougados et al., 1999 Medium</td>
<td>RCT N=209 1 yr 5-yr followup</td>
<td>SSZ (2-3 g/day) + MTX (7.5 to 15 mg/week) vs. SSZ vs. MTX</td>
<td>Overall discontinuation: 29.2%, 30.9%, 21.7% Discontinuation due to adverse events: 12.5%, 14.7%, 10.1% Serious adverse events: 1.0%, 0.0%, 2.0% Specific adverse events: Nausea: 49.0%, 32.0%, 23.0%, p=0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Haagsma 1997 Medium</td>
<td>RCT N=105 1 yr</td>
<td>SSZ (1-3 g/day) vs. MTX (7.5-15 mg/week) vs. MTX + SSZ</td>
<td>Overall discontinuation: 35.3%, 5.7%, 16.7% Discontinuation due to adverse events: 26.5%, 5.7%, 13.9% Serious adverse events: 8.8%, 0.0%, 0.0% Specific adverse events: Nausea: 29.4%, 25.7%, 63.9% Upper respiratory infection: 17.6%, 20.0%, 27.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr &amp; Risk of Bias Ratings</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>Nijmegen RA Inception 2009</td>
<td>Observational N=230 1 yr</td>
<td>(SSZ failures) Switch from SSZ to MTX (7.5 mg-30 mg/week) vs. MTX and continue SSZ (750-3,000 mg/day)</td>
<td>Overall discontinuation: 33.9% vs. 50.0%, p=0.013 Discontinuation due to adverse events: 18.5%, 11.3% Serious adverse events: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td>High for 12 months</td>
<td>Specific adverse events: NR</td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>COBRA 1997, COBRA 2002</td>
<td>RCT N=155 5 yrs</td>
<td>PNL (60 mg tapered over 28 weeks) + MTX (7.5 mg/week stopped after 40 weeks) + SSZ (2,000 mg/day) vs. SSZ</td>
<td>Overall discontinuation: 8.0% vs. 29.1%, p=0.0008 Discontinuation due to adverse events: 2.6% vs. 7.6% Serious adverse events: 2.6% vs. 7.6% Specific adverse events: GI complaints: 14.5% vs. 12.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>COBRA Light, 2014</td>
<td>RCT N=164 1 yr</td>
<td>PNL (60 mg tapered to 7.5 mg/day) + MTX 7.5 mg/week) + SSZ (2 g/day) vs. PNL (30 mg/d tapered to 7 mg/day) + MTX (25 mg/week)</td>
<td>Overall discontinuation: 3.7% vs. 4.9% Discontinuation due to adverse events: NR Serious adverse events: 11.1% vs. 19.8% Specific adverse events: Leukopenia: 1.0% vs. 4.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>FIN-RACO 1999, FIN-RACO 2010, FIN-RACO 2013, FIN-RACO 2004, FIN-RACO 2004, FIN-RACO 2010</td>
<td>RCT N=199 2 yrs 5-yr followup</td>
<td>MTX (7.5-10 mg/week) + HCQ (300 mg/day) + SSZ (2 g/day) + PNL (5-10 mg/day) vs. DMARD (SSZ could be changed to MTX if adverse event or lack of response)</td>
<td>Overall discontinuation: 10.3% vs. 7.1% Discontinuation due to adverse events: 23.7% vs. 22.4% Serious adverse events: 3.1%, 5.1% Specific adverse events: Elevated liver enzymes (AAT and AP &gt; 2x normal): 11.3% vs. 23.5%, p=0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARD Combination Therapy vs. csDMARD Monotherapy</td>
<td>tREACH 2013, tREACH 2014, tREACH 2016</td>
<td>RCT N=515 1 yr</td>
<td>MTX (25 mg/week) + SSZ (2 g/day) + HCQ (400 mg/day) + GCs intramuscularly vs. MTX + SSZ + HCQ + GC oral taper (15 mg/day tapers off at 10 weeks) vs. MTX + GC oral taper</td>
<td>Overall discontinuation: 15% vs. 9.7% vs. 10.3% Discontinuation due to adverse events: 1.1%, 0.0%, 2.1% Serious adverse events: 5.0%, 11.0%, 10.0% Specific adverse events: Headache: 11.0% vs. 14.0% vs. 13.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Ratings</td>
<td>Study Design N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>----------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>TNF Biologic + csDMARD vs. TNF biologic</td>
<td>PREMIER 2006,13 2008,103 2010,149 2010,115 2012,116 2013,117 2014,118 2015119,c</td>
<td>Medium</td>
<td>RCT N=799</td>
<td>2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (20 mg/week) vs. ADA vs. MTX</td>
<td>Overall discontinuation: 24.3% vs. 39.1% vs. 34.2%, p&lt;0.001 Discontinuation due to adverse events: 11.9% vs. 9.5% vs. 7.4%, p=0.21 Serious adverse events: 18.5%, 21.1%, 15.9%, p=0.192 Specific adverse events: Higher serious infections (n per 100 pt-years) in ADA + MTX vs. ADA: 2.9, 0.7, p&lt;0.05</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic</td>
<td>AVERT, 20157 a d</td>
<td>Medium</td>
<td>RCT N=351</td>
<td>2 yrs</td>
<td>ABA (125 mg/week) + MTX (7.5-20 mg/week) vs. ABA vs. MTX</td>
<td>Overall discontinuation: 13.4%, 21.6%, 17.2% Discontinuation due to adverse events: 1.7%, 4.3%, 2.6% Serious adverse events: 6.7%, 12.1%, 7.8% Specific adverse events: Serious infection: 0.8% vs. 3.4% vs. 0%</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic</td>
<td>FUNCTION 201612 a d</td>
<td>Medium</td>
<td>RCT N=1,162</td>
<td>1 yr</td>
<td>TCZ (4 mg/kg monthly) + MTX (20 mg/week) vs. TCZ (8 mg/kg monthly) + MTX vs. TCZ vs. MTX</td>
<td>Overall discontinuation: 20.3%, 22%, 19.2%, 21.8% Discontinuation due to adverse events: 12.1%, 20.3%, 11.6%, 7.4% Serious adverse events: 10%, 10.7%, 8.6%, 8.5% Specific adverse events: NR</td>
</tr>
<tr>
<td>Non-TNF Biologic + csDMARD vs. Non-TNF Biologic</td>
<td>U-Act-Early 201613 a d</td>
<td>Medium</td>
<td>RCT N=317</td>
<td>2 yrs</td>
<td>TCZ (8 mg/kg monthly) + MTX (10-30 mg/week) vs. TCZ vs. MTX</td>
<td>Overall discontinuation: 26.4%, 21.4%, 27.8% Discontinuation due to adverse events: 8.5%, 9.7%, 7.4%, p=0.82 Serious adverse events: 16%, 18.4%, 12%, p=0.44 Specific adverse events: NR</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Risk of Bias Ratings</td>
<td>Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------</td>
<td>-----------------------------</td>
<td>--------</td>
<td>---</td>
<td>---------</td>
<td>------------------</td>
</tr>
<tr>
<td>csDMARDs vs. tsDMARDs</td>
<td>Conaghan 2016&lt;sup&gt;39&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>N=108</td>
<td>1 yr</td>
<td>TOF (20 mg/day) + MTX (10-20 mg/week) vs. TOF vs. MTX</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>HIT HARD 2013&lt;sup&gt;44&lt;/sup&gt;</td>
<td>Medium (DAS, ACR)</td>
<td>RCT</td>
<td>N=172</td>
<td>48 weeks</td>
<td>ADA (40 mg biweekly x 24 weeks) + MTX (15 mg/week) vs. MTX</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>HOPEFUL 1 2014&lt;sup&gt;45,150&lt;/sup&gt;</td>
<td>Medium</td>
<td>RCT</td>
<td>N=334</td>
<td>52 weeks</td>
<td>ADA (40 mg biweekly) + MTX (6-8 mg/week) vs. MTX</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>OPTIMA 2013&lt;sup&gt;47&lt;/sup&gt;, 2014&lt;sup&gt;151&lt;/sup&gt;, 2016&lt;sup&gt;152 a&lt;/sup&gt;</td>
<td>Low</td>
<td>RCT</td>
<td>N=1,032</td>
<td>78 weeks</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/week) vs. MTX</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>PREMIER 2006,15, 2008,103 2010,109 2010,115 2012,116 2013,117 2014,118 2015119 c</td>
<td>RCT</td>
<td>N=799</td>
<td>2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (20 mg/week) vs. ADA vs. MTX</td>
<td>Overall discontinuation: 24.3% vs. 39.1% vs. 34.2%, p&lt;0.001 Discontinuation due to adverse events: 11.9% vs. 9.5% vs. 7.4%, p=0.21 Serious adverse events: 18.5%, 21.1%, 15.9%, p=0.192 Specific adverse events: Higher rates of serious infections (n per 100 pt-years) in ADA + MTX vs. ADA: 2.9, 0.7, p&lt;0.05</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>PROWD 200819, 201612</td>
<td>RCT</td>
<td>N=148</td>
<td>56 weeks</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/week) vs. MTX</td>
<td>Overall discontinuation: 25.0% vs. 37.0% Discontinuation due to adverse events: 8.0% vs. 11.0% Serious adverse events: 17.3% vs. 15.1% Specific adverse events: Abdominal pain: 1.4% vs. 0.0% Nausea: 21.3% vs. 32.9% Diarrhea: 10.7% vs. 8.2% Headache: 10.7% vs. 6.8%</td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>C-OPERA 201613 a</td>
<td>RCT</td>
<td>N=316</td>
<td>2 yrs</td>
<td>CZP (400 mg biweekly x 4 weeks, then 200 mg biweekly) + MTX (8-12 mg/week) vs. MTX</td>
<td>Overall discontinuation: 53.5% vs. 63.7% Discontinuation due to adverse events: 6.3% vs. 3.8% Serious adverse events: 10.7% vs. 11.5% Specific adverse events: Nausea: 27.0% vs. 24.2% Injection site reaction: 3.1% vs. 1.3% Interstitial Lung disease: 4.4% vs. 0.6% Hepatic disorders: 42.8% vs. 44.6%</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr Risk of Bias Ratings</td>
<td>Study Design Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>C-EARLY 2017&lt;sup&gt;18, 39&lt;/sup&gt; Medium</td>
<td>RCT N=879 52 weeks Aggressive RA</td>
<td>CZP (400 mg biweekly) + MTX (10-25 mg/wk) vs. MTX</td>
<td>Overall discontinuation: 24.2% vs. 34.7% Discontinuation due to adverse events: 7.7 vs. 7.8%, p=NS, NR Serious adverse events: 10.6% vs. 9.2%, p=NS, NR Specific adverse events: Nausea: 12.6% vs. 10.1% Upper respiratory tract infection: 10.9% vs. 5.1% Urinary tract infection: 7.3% vs. 7.4% Headache: 6.8% vs. 3.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>COMET 2008&lt;sup&gt;12, 108, 109, 154-156 a&lt;/sup&gt; Medium</td>
<td>RCT N=542 2 yrs</td>
<td>ETN (50 mg/week) + MTX (7.5 mg/week) vs. MTX</td>
<td>Overall discontinuation: 19.3% vs. 29.5% Discontinuation due to adverse events: 10.2% vs. 12.7% Serious adverse events: 12.0% vs. 12.7% Specific adverse events: Malignancy: 1.5% vs. 1.5% Upper respiratory infection: 45.0% vs. 44.0% Nausea: 53.0% vs. 50.0% Infusion/injection site reactions: 1.0% vs. 2.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Enbrel ERA 2000&lt;sup&gt;14, 110-112 a&lt;/sup&gt; Medium</td>
<td>RCT N=632 1 yr (1-yr open-label extension)</td>
<td>ETN (25 mg twice weekly) vs. MTX (20 mg/week)</td>
<td>Overall discontinuation: 25.6% vs. 40.5% Discontinuation due to adverse events: 7.3% vs. 12.4% Serious adverse events: 12.0% vs. 12.0% Specific adverse events: Injection site reaction: 39.0% vs. 9.0%, p&lt;0.05 Nausea: 20.0% vs. 31.0%, p&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Marcora et al., 2006</td>
<td>RCT</td>
<td>26</td>
<td>24 weeks</td>
<td>ETN (25 mg twice weekly) vs. MTX (7.5-15 mg/week)</td>
<td>Overall discontinuation: 0.0% vs. 0.0%</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to adverse events: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Serious adverse events: 0.0% vs. 0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific adverse events: Injection site reaction: 8.3% vs. 0.0%</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>ASPIRE 2004</td>
<td>RCT</td>
<td>1,049</td>
<td>54 weeks</td>
<td>IFX (3 mg/kg/8 weeks) + MTX (20 mg/week) vs. IFX (6 mg/kg/8 weeks) + MTX vs. MTX</td>
<td>Overall discontinuation: 21.4%, 23.8%, 25.5%</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to adverse events: 9.5%, 9.6%, 3.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Serious adverse events: 11.0%, 14.0%, 14.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific adverse events: Infusion or injection site reaction: 21.0%, 15.0%, 7.0% TB: 0.8%, 0.3%, 0.0% Serious infection: 5.6%, 5.0%, 2.1%, p=0.02</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Quinn et al., 2005</td>
<td>RCT</td>
<td>20</td>
<td>2 yrs</td>
<td>IFX 3 mg/kg 0, 2, 6, and every 8 weeks) + MTX (7.5-25 mg/wk) vs. MTX (7.5-25 mg/week)</td>
<td>Overall discontinuation: NR</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to adverse events: 5.0% overall</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Serious adverse events: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific adverse events:</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Monotherapy</td>
<td>Durez et al., 2007</td>
<td>RCT</td>
<td>44</td>
<td>1 yr</td>
<td>IFX (3 mg/kg at weeks 0, 2, 6 until 46 weeks) + MTX (7.5-20 mg/wk) vs. MTX</td>
<td>Overall discontinuation: 6.7% vs. 14.3%</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to adverse events: 6.7% vs. 0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific adverse events: Benign infection: 80.0% vs. 93.3% Mild hepatotoxicity: 14.3% vs. 33.5%</td>
</tr>
<tr>
<td>TNF Biologic vs. csDMARD Combination Therapy</td>
<td>IMPROVE D, 2013</td>
<td>RCT</td>
<td>161</td>
<td>2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (25 mg/wk) vs. MTX + PRED (7.5 mg/day) + HCQ (400 mg/day) + SSZ (2 g/day)</td>
<td>Overall discontinuation: NR</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to adverse events: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific adverse events: Increase liver enzymes: 8.4% vs. 4.0%</td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Ratings</td>
<td>Study Design N Duration</td>
<td>Comparison (Dose)</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Combination Therapy</strong></td>
<td>SWEFOT, 2013</td>
<td>Medium</td>
<td>N=258 1 yr</td>
<td>IFX (3 mg/kg at 0,2,6 weeks then biweekly) + MTX (20 mg/wk) vs. MTX + SSZ (2 g/day) + HCQ (400 mg/day)</td>
<td>Overall discontinuation: 31.5% vs. 18.0%, p = 0.014 Discontinuation due to adverse events: 10.8% vs. 7.8% Specific adverse events: GI symptoms (not specified): 11.5% vs. 0.7% Skin and allergic reactions: 2.3% vs. 8.5%</td>
<td></td>
</tr>
<tr>
<td><strong>TNF Biologic vs. csDMARD Combination Therapy</strong></td>
<td>NEO-RACo, 2013</td>
<td>Low</td>
<td>N=99 2 yrs</td>
<td>IFX (3 mg/kg) + FIN-RACo (MTX 25 mg/week) + SSZ 2 g/day) + HCQ (35 mg/kg/week) + PRED (7.5 mg/day) for 26 weeks vs. FIN-RACo</td>
<td>Overall discontinuation: 8% vs. 8.2% Discontinuation due to adverse events: 2.0% vs. 0.0% Serious adverse events: 6.0% vs. 8.0% Specific adverse events: GI: 56.0% vs. 61.0% Respiratory: 56% vs. 67.0% Elevated liver enzymes: 12.0% vs. 16.0% No significant differences between arms overall</td>
<td></td>
</tr>
<tr>
<td><strong>Non-TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>AGREE, 2009</td>
<td>Low</td>
<td>N=509 2 yrs</td>
<td>ABA (10 mg/kg) + MTX (7.5 mg/week) vs. MTX</td>
<td>Overall discontinuation: 9.4% vs. 10.3% Discontinuation due to adverse events: 3.1% vs. 4.3% Serious adverse events: 7.8% vs. 7.9% Specific adverse events: Upper respiratory infection: 10.2% vs. 10.3%</td>
<td></td>
</tr>
<tr>
<td><strong>Non-TNF Biologic vs. csDMARD Monotherapy</strong></td>
<td>AVERT, 2015</td>
<td>Medium</td>
<td>N=351 2 yrs</td>
<td>ABA (125 mg/week) + MTX (7.5-20 mg/week) vs. ABA vs. MTX</td>
<td>Overall discontinuation: 13.4%, 21.6%, 17.2% Discontinuation due to adverse events: 1.7%, 4.3%, 2.6% Serious adverse events: 6.7%, 12.1%, 7.8% Specific adverse events: Serious infection: 0.8% vs. 3.4% vs. 0%</td>
<td></td>
</tr>
<tr>
<td>Drug Therapy Comparison Category</td>
<td>Study, Yr</td>
<td>Risk of Bias Ratings</td>
<td>Study Design</td>
<td>N</td>
<td>Duration</td>
<td>Comparison (Dose)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
<td>----------------------</td>
<td>--------------</td>
<td>---</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>FUNCTION 2016</td>
<td>Medium</td>
<td>RCT</td>
<td>1,162</td>
<td>1 yr</td>
<td>TCZ (4 mg/kg monthly) + MTX (20 mg/week) vs. TCZ (8 mg/kg monthly) + MTX vs. TCZ vs. MTX</td>
</tr>
<tr>
<td></td>
<td>IMAGE, 2012</td>
<td>Low</td>
<td>RCT</td>
<td>755</td>
<td>2 yrs</td>
<td>RIT (1 g days 1 and 15) + MTX (7.5-30 mg/week) vs. RIT (500 mg days 1 and 15) + MTX vs. MTX</td>
</tr>
<tr>
<td></td>
<td>U-Act-Early 2016</td>
<td>Medium</td>
<td>RCT</td>
<td>317</td>
<td>2 yrs</td>
<td>TCZ (8 mg/kg monthly) + MTX (10-30 mg/week) vs. TCZ vs. MTX</td>
</tr>
<tr>
<td>TNF vs. Non-TNF</td>
<td>ORBIT, 2016</td>
<td>High</td>
<td>RCT</td>
<td>329</td>
<td>1 yr</td>
<td>RIT (1 g on days 1 and 15 and after 26 if persistent disease activity) vs. ADA (40 mg biweekly) or ETN (50 mg/week)</td>
</tr>
</tbody>
</table>

Specific adverse events:
- Non-TNF Biologic vs. csDMARD Monotherapy Specific adverse events: NR
- Non-TNF Biologic vs. csDMARD Monotherapy Specific adverse events: Infusion-related reaction: 18.4% vs. 14.1% vs. 12.4%
- Non-TNF Biologic vs. csDMARD Monotherapy Specific adverse events: Infections: 53.5% vs. 70.9%
- Non-TNF Biologic vs. csDMARD Monotherapy Specific adverse events: Injection site reactions less with RIT p=0.003
<table>
<thead>
<tr>
<th>Drug Therapy Comparison Category</th>
<th>Study, Yr</th>
<th>Study Design</th>
<th>Comparison (Dose)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combination and Therapy Strategies</td>
<td>BeSt, 2005&lt;sup&gt;79,91&lt;/sup&gt;</td>
<td>RCT N=508 12 months plus 10-yr followup</td>
<td>DAS-driven treatment; 1: sequential monotherapy starting with MTX (15 mg/week) vs. 2: stepped-up combination therapy: MTX, then SSZ, then HCQ, then PRED vs. 3: combination with tapered high-dose PRED (60 mg/d to 7.5 mg/day) vs. 4: combination MTX (25-30 mg/week) with IFX (3 mg/kg every 8 weeks, per DAS, could be titrated to 10 mg/kg)</td>
<td>5 yrs Overall discontinuation: 12.0%, 22.0%, 15.0%, 9.0%; 2 vs. 4, p=0.05 Discontinuation due to adverse events: NR Serious adverse events: 33.0%, 28.0%, 28.0%, 31.0%, p=0.76 Specific adverse events: NR 10 yrs No significant differences in serious adverse events (SAE per 100 pt yrs) 13.2, 10.9, 12.1, 13.4, p=0.47</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>TEAR, 2012&lt;sup&gt;20,159&lt;/sup&gt;</td>
<td>RCT N=755 2 yrs</td>
<td>1: immediate MTX plus ETN vs. 2: immediate MTX plus SSZ plus HCQ vs. 3: step-up MTX to MTX + ETN vs. 4: step-up MTX to MTX + SSZ + HCQ</td>
<td>Overall discontinuation: 42.4%, 34.8%, 39.5%, 34.9% Discontinuation due to adverse events: 1&amp;2: 1.9%, 3&amp;4: 1.3% Serious adverse events: 13.6%, 14.3%, 12.9%, 12.5%, p=0.94 Specific adverse events: NR</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>GUEPARD 2009&lt;sup&gt;22&lt;/sup&gt;</td>
<td>RCT N=65 1 yr</td>
<td>1: ADA 40 mg every 2 weeks plus MTX; treatment adjusted every 3 months to achieve DAS28 &lt;3.2 2: MTX (max 20 mg/wk)</td>
<td>Overall discontinuation: 15.2% vs. 9.4% Discontinuation due to adverse events: NR Serious adverse events: 15.2% vs. 15.6% Specific adverse events: NR</td>
</tr>
<tr>
<td>Combination and Therapy Strategies</td>
<td>OPERA 2017&lt;sup&gt;160-163&lt;/sup&gt;</td>
<td>RCT N=180 2 yrs</td>
<td>ADA (40 mg biweekly) + MTX (7.5-20 mg/week) vs. MTX</td>
<td>Overall discontinuation: 10.1% vs. 16.5% Discontinuation due to adverse events: 2.2% vs. 1.1% Serious adverse events: 4% vs. 11% Specific adverse events: Bronchitis: 1.1% vs. 1.1% Leukopenia: 0% vs. 1.1%</td>
</tr>
</tbody>
</table>
### Drug Therapy Comparison Category

<table>
<thead>
<tr>
<th>Study, Yr</th>
<th>Study Design</th>
<th>Comparison (Dose)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bili et al., 2014</td>
<td>Observational</td>
<td>1: TNFα inhibitors alone or in combination with MTX</td>
<td>Overall discontinuation: NR</td>
</tr>
<tr>
<td>High</td>
<td>N=2,101 10 yrs</td>
<td>2: MTX alone or in combination with other nonbiologic DMARDs</td>
<td>Discontinuation due to adverse events: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3: Non-MTX, nonbiologic DMARDs</td>
<td>Serious adverse events: NR</td>
</tr>
<tr>
<td>ERAN Inception Cohort, 2013</td>
<td>Observational</td>
<td>1: Initial SSZ</td>
<td>Overall discontinuation: NR</td>
</tr>
<tr>
<td>High</td>
<td>N=766 2 yrs</td>
<td>2: Initial MTX</td>
<td>Discontinuation due to adverse events: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3: MTX + SSZ+ HCQ</td>
<td>Serious adverse events: NR</td>
</tr>
</tbody>
</table>

* Included in network meta-analysis (NWMA)
* This study evaluates comparisons in both the High-Dose Corticosteroid and TNF Biologic vs. csDMARD monotherapy categories.
* This study evaluates comparisons in both the csDMARD vs. TNF Biologic and TNF Biologic vs. csDMARD monotherapy categories.
* These studies evaluate comparisons in both the csDMARD vs. Non-TNF Biologic and Non-TNF Biologic vs. csDMARD monotherapy categories.

AAT = alanine aminotransferase; ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; ALT = alanine transaminase; AP = alkaline phosphatase; AST = aspartate aminotransferase; csDMARD = conventional synthetic disease-modifying antirheumatic drug; CZP = certolizumab pegol; DAS = Disease Activity Score (based on 44 joints); DMARD = disease modifying antirheumatic drug; ETN = etanercept; g = grams; GC = glucocorticoid; GI = gastrointestinal; HCQ = hydroxychloroquine; IFX = infliximab; kg = kilograms; LEF = leflunomide; methyl-PNL = methylprednisolone; mg = milligram; mg/d = milligrams per day; MTX = methotrexate; N = number; NR = not reported; PNL = prednisolone; PRED = prednisone; pt-years = patient-years; RCT = randomized controlled trial; RIT = rituximab; SHS = Sharp/van der Heijde Score; SSZ = sulfasalazine; TB = tuberculosis; TCZ = tocilizumab; TNF = tumor necrosis factor; TOF = tofacitinib; ULN = upper limit of normal; vs. = versus; wk = week.

In the detailed synthesis below, we report on these outcomes separately for RCTs and observational studies. The evidence primarily includes RCTs. The results of our NWMA (network diagrams and forest plots) are presented below in figures accompanying the results for specific drug comparisons.

Because of the dearth of trials directly comparing interventions of interest, we employed NWMA. For KQ 3, we conducted NWMA on the following outcomes: all discontinuations (unintended for any reason such as an adverse event, side effect, lack of effectiveness or any other reason to drop out of a study) (16 trials) and discontinuations due to adverse events. For NWMA, we focused on a time period around 1 year (52 to 56 weeks) because data were more comprehensive for this time period than for other ones. For other time points, data were insufficient for NWMA, or the clinical heterogeneity across trials was too high to derive meaningful estimates from NWMA. We detected no significant differences between the consistency and inconsistency models for these two outcomes (see Appendix G for details). Therefore, we report estimates based on the consistency models. We present results of NWMA...
for all discontinuations and discontinuations because of adverse events within each comparison section below.

Figure 19 depicts the network diagram for both outcomes, and Table 11 lists the studies we used in each NWMA. The network structure is mostly “star-shaped,” indicating a dearth of head-to-head studies directly comparing interventions. Most effect estimates, therefore, were derived from indirect comparisons relative to MTX rather than mixed treatment comparisons. Our NWMA for all discontinuations and for discontinuations attributable to adverse events were reported below. Confidence intervals for the NWMA for discontinuations and discontinuations due to adverse events were wide and should be interpreted with caution.

**Figure 19. Network diagram for network meta-analysis: All discontinuations and discontinuations due to adverse events**

MTX = methotrexate; N = number of patients.
Table 11. Studies included in KQ 3 network meta-analysis

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>Study Name</th>
<th>Overall D/C</th>
<th>D/C due to AE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA + MTX vs. MTX</td>
<td>AGREE, 2009,2015; 2011,129,130</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ABA + MTX vs. ABA vs. MTX</td>
<td>AVERT, 2015</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ADA + MTX vs. MTX</td>
<td>PROWD, 2008,16, 2016</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>C-EARLY, 2017; 38, 39</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>C-OPERA, 2016,13, 2017</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IFX + MTX vs. Methyl-PNL + MTX vs. MTX</td>
<td>Durez et al., 2007</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IFX + MTX vs. MTX</td>
<td>Quinn et al., 2005</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SSZ + MTX vs. SSZ vs. MTX</td>
<td>Dougados et al., 1999;31, Maillefert et al., 2003</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SSZ + MTX vs. SSZ vs. MTX</td>
<td>Haagsma et al., 1997</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ vs. MTX</td>
<td>FUNCTION, 2016</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ vs. MTX</td>
<td>U-Act-Early, 2016</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

a All data used in NWMA were measured at the 1-year followup time point.

ABA = abatacept; ACR50 = American College of Rheumatology 50% improvement; ADA = adalimumab; AE = adverse event; AGREE = Abatacept trial to Gauge Remission and joint damage progression in methotrexate-naïve patients with Early Erosive rheumatoid arthritis; ASPIRE = Active-controlled Study of Patients receiving Infliximab for the treatment of Rheumatoid arthritis of Early onset trial; AVERT = Assessing Very Early Rheumatoid arthritis Treatment trial; C-EARLY = trial whose acronym not described; C-OPERA = Certolizumab-Optimal Prevention of joint damage for Early RA trial; COMET = Combination of Methotrexate and Etanercept in Active Early Rheumatoid Arthritis trial; CZP = certolizumab pegol; D/C = discontinuation; Enbrel ERA = Enbrel Early RA trial; ETN = etanercept; FUNCTION = trial whose acronym not described; IFX = infliximab; KQ = Key Question; methyl-PNL = methylprednisolone; MTX = methotrexate; NA = not applicable; NWMA = network meta-analysis; RA = rheumatoid arthritis; SSZ = sulfasalazine; TCZ = tocilizumab; U-Act-Early = Trial whose acronym not described; vs. = versus.

Corticosteroids

Corticosteroids Versus csDMARDs

Five trials examined overall risk of harms, discontinuation, adherence, serious adverse events, and specific adverse events (Table 10). Many of the csDMARD investigations involved a corticosteroid plus a csDMARD (majority with MTX) compared with csDMARD monotherapy. Corticosteroids and csDMARDs did not differ significantly in serious adverse events (moderate SOE) or discontinuations attributable to adverse events (low SOE). Over 2 years, discontinuation rates in the combination corticosteroid plus csDMARD arm ranged from 8.2 percent to 47.0 percent; in the csDMARD arm, the rates ranged from 10.9 percent to 29.8 percent. Overall, no significant differences were found in discontinuations attributed to adverse events and serious adverse events. The CAMERA-II trial reported nausea significantly
less in the PRED plus MTX arm than in the MTX monotherapy arm (19.6% vs. 36.1%, p=0.006).\textsuperscript{94} Additionally, elevated transaminases occurred less often in the PRED plus MTX arm.\textsuperscript{94} These could be chance findings because we could not find consistent findings in the other studies. Occurrences of infection did not differ significantly in either the CAMERA-II or the CARDERA trials.\textsuperscript{93, 94}

**High-Dose Corticosteroids**

Two trials compared the combination of IFX plus MTX with high-dose methyl-PNL and MTX.\textsuperscript{18, 96} Overall, the SOE was insufficient for discontinuations because of adverse events and serious adverse events. The IDEA trial (N=112)\textsuperscript{96} lasted for 26 weeks, and then patients were converted to open-label treatment for an additional 50 weeks. The investigators reported no appreciable differences in overall discontinuation and discontinuation attributable to adverse events (5.5% vs. 1.8%, p=NR). However, reported serious adverse events were 36.4 percent in the MTX plus IFX group and 15.8 percent in the high-dose methyl-PNL plus MTX group (p=NR). These included admissions for surgical procedures unrelated to RA or to study treatment and serious infections. Upper respiratory infections were similar (3.6% vs. 1.8%, p=NR). In the second smaller trial (N=44),\textsuperscript{18} overall discontinuations were 6.7 percent for IFX plus MTX and methyl-PNL plus MTX and numerically higher (14.3%) for MTX monotherapy (p=NR). Only one person randomized to the IFX plus MTX group experienced a serious adverse event (MTX-related pneumonia at week 30). Other side effects were equally distributed between the groups (benign infection and mild hepatotoxicity).

**Single-Arm Study: Corticosteroids Only**

One single-arm observational cohort study (N=12,656) examined patients in the Swedish Rheumatology Quality Register with incident RA, matched them to 10 population comparator patients, and followed them over 12 years for lymphoma risk.\textsuperscript{76} After adjustment for age, sex, and inflammatory activity during the first year of RA diagnosis, corticosteroid use was associated with a reduced risk of lymphoma (RR, 0.5; 95% CI, 0.3 to 0.9).

**csDMARDs**

**csDMARDs Versus csDMARDs**

**csDMARD Monotherapy Versus csDMARD Monotherapy**

One trial\textsuperscript{27} compared MTX plus prednisolone (PNL) with SSZ plus PNL, and one observational study\textsuperscript{28} compared MTX with SSZ. In both studies, overall discontinuation rates and discontinuation rates attributable to adverse events were higher for SSZ than for MTX. Overall, the SOE based on either study was insufficient for discontinuations because of adverse events and serious adverse events. Our NWMA supported this finding with higher overall discontinuations for SSZ compared with MTX (RR, 1.83; 95% CI, 1.06 to 3.16) (Figure 20). However, differences in discontinuations due to adverse events were not significant (Figure 21).

In the observational study (N=1,102), the specific adverse events were mixed depending on the drug group.\textsuperscript{28} The SSZ group experienced significantly higher abdominal pain (8.0% vs. 4.0%, p<0.03) and rash (9.1% vs. 2.7%, p<0.001). The MTX group, however, experienced significantly higher rates of infection (34.1% vs. 20%, p<0.001) and nausea (18.9% vs. 13.1%, p<0.07).
csDMARD Combination Therapy Versus csDMARD Monotherapy

csDMARD combination therapy compared with csDMARD monotherapy did not differ significantly in serious adverse events (low SOE). Six trials compared SSZ plus MTX with csDMARD monotherapy (MTX or SSZ).4, 21-24, 105 Overall discontinuations were mixed. The majority of the trials found no significant differences between SSZ plus MTX groups and csDMARD-only groups. In one 5-year trial (N=155), however, discontinuation rates were higher in the SSZ monotherapy arm than in the MTX plus SSZ (29.1% vs. 8.0%, p=0.0008).24

In addition, one observational study (N=230) found higher rates of overall discontinuation in the MTX plus SSZ group than in the MTX-only group (50.0% vs. 33.9%, p=0.013).26 However, no significant differences occurred in discontinuations due to adverse events (insufficient SOE).

csDMARDs Versus Biologics

TNF Biologic: MTX Plus TNF Biologic Versus Monotherapy With Either MTX or TNF Biologic

Combining a csDMARD with a TNF biologic did not differ significantly in serious adverse events (moderate SOE) or discontinuations attributable to adverse events compared with
csDMARD monotherapy (moderate SOE). The PREMIER trial (N=799) examined combination therapy with MTX plus ADA compared with monotherapy with either MTX or ADA in patients with early aggressive RA.\textsuperscript{15} After 2 years, the MTX plus ADA arm had lower discontinuation rates than either the ADA or MTX monotherapy arm (24.3\% vs. 39.1\% vs. 34.2\%, p<0.001). Neither discontinuations attributable to adverse events (11.9\% vs. 9.5\% vs. 7.4\%, \textit{p}=0.21) nor the proportion of serious adverse events differed significantly by group (18.5\% vs. 21.1\% vs. 15.9\%, \textit{p}=0.19). Our NWMA examined ETN plus MTX versus ETN and found no significant differences in all discontinuations (Figure 22) or discontinuations due to adverse events (Figure 23).

**Figure 22. Forest plot for network meta-analysis of all discontinuations: TNF + MTX versus TNF**

![Forest plot for network meta-analysis of all discontinuations: TNF + MTX versus TNF](image1)

95\% CI = 95\% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

**Figure 23. Forest plot for network meta-analysis of discontinuations due to adverse events: TNF + MTX versus TNF**

![Forest plot for network meta-analysis of discontinuations due to adverse events: TNF + MTX versus TNF](image2)

95\% CI = 95\% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

**Non-TNF Biologic: MTX Plus Non-TNF Biologic Versus Monotherapy With Either MTX or Non-TNF Biologic**

One trial compared the combination of ABA plus MTX with either ABA or MTX monotherapy: the AVERT study (N=351).\textsuperscript{7} It found no significant differences in overall discontinuation rates, discontinuation attributable to adverse events, or serious adverse events.

Two RCTs examined discontinuation rates for patients receiving combination therapy with TCZ plus MTX and patients receiving either MTX or TCZ monotherapy: the FUNCTION 2-year trial (N=1,162)\textsuperscript{32,134} and the U-Act-Early 2-year trial (N=317).\textsuperscript{33} Overall discontinuation rates
and discontinuation attributable to either adverse events (U-Act-Early: 8.5% vs. 9.7% vs. 7.4%, \(p=0.82\)) or serious adverse events (U-Act-Early: 16.0% vs. 18.4% vs. 12.0%, \(p=0.44\)) did not differ across these groups (moderate SOE).

The NWMA similarly found no significant differences in overall discontinuations or discontinuations attributable to adverse events for TCZ monotherapy compared with TCZ plus MTX. Figure 24 presents findings for all discontinuations and Figure 25 for discontinuations attributable to adverse events; in both cases, results are reported as RRs with 95% CIs. NWMA also examined ABA plus MTX and found no significant differences in overall discontinuations but fewer discontinuations due to adverse events for ABA plus MTX than ABA monotherapy (RR, 0.34; 95% CI, 0.18 to 0.64).

**Figure 24. Forest plot for network meta-analysis of all discontinuations: Non-TNF + MTX versus non-TNF**

<table>
<thead>
<tr>
<th></th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. Abatacept</td>
<td>0.66 (0.39, 1.11)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Tocilizumab</td>
<td>1.20 (0.88, 1.63)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
csDMARDs versus tsDMARDs: MTX Plus tsDMARD Versus Either MTX or tsDMARD

One RCT (N=109) compared the combination of TOF plus MTX with monotherapy (TOF or MTX) over 12 months in patients with early active RA. Overall discontinuation rates were 21.4 percent for the combination therapy group, 43.2 percent for TOF monotherapy, and 25.0 percent for MTX monotherapy. The groups did not have any significant differences for discontinuations attributable to adverse events (TOF monotherapy, 5.6%; MTX monotherapy, 13.5%; TOF plus MTX therapy, 11.1%). Additionally, no differences in serious adverse events were reported for patients receiving TOF monotherapy (2.8%), MTX monotherapy (5.4%), or TOF plus MTX therapy (5.6%) (insufficient SOE).

Single-Arm Studies: csDMARDs Only

Four single-arm observational studies examined various approaches to using csDMARDs. One involved a three-csDMARD regimen (MTX plus SSZ plus either HCQ or LEF); another study focused only on LEF, a third on MTX exposure or TNFi (i.e., TNF biologic exposure), and a fourth only on MTX. SSZ was the most common drug removed from triple therapy because of adverse events (49.0%) over 70 weeks, followed by MTX (29.0%) and HCQ (13.0%). A 15-year retrospective observational study examined exposure to RA drugs in the first year (csDMARDs, corticosteroids, biologics) and subsequent lymphoma diagnosis and found no increased lymphoma risk in patients exposed to MTX (RR, 0.9; 95% CI, 0.8 to 1.0) in the first year of diagnosis compared with RA patients. In a 1-year prospective study of LEF, overall discontinuation was 11.1 percent. In a cohort of patients with early RA taking MTX, 50 percent discontinued after 10.9 years (reasons for discontinuation not described).
Biologics

TNF Biologics

TNF Biologic Versus csDMARD Monotherapy

Neither serious adverse events nor discontinuations attributable to adverse events differed significantly between the TNF biologics (ADA, CZP, ETN, IFX) in combination with MTX versus MTX monotherapy (low SOE). In NWMA, TNF biologics (ADA, CZP, ETN, IFX) plus MTX had lower overall discontinuations than the csDMARD SSZ (range of RR, 0.35 to 0.48 [95% CI, 0.18 to 0.89]); only IFX plus MTX had higher discontinuation resulting from adverse events (RR, 3.03; 95% CI, 1.56 to 5.90) (Figure 26 and Figure 27, respectively).

Adalimumab. Five RCTs examined the combination of ADA plus MTX with MTX monotherapy over 26 weeks to 2 years. In general, no significant differences were observed for discontinuations due to adverse events or serious adverse events (low SOE). In NWMA, there were no differences in overall discontinuations or discontinuations due to adverse events (Figure 26 and Figure 27, respectively).

Figure 26. Forest plot for network meta-analysis of all discontinuations: TNF + MTX versus csDMARD

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adalimumab+MTX vs. MTX</td>
<td>0.66 (0.43, 1.00)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. MTX</td>
<td>0.64 (0.52, 0.78)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. MTX</td>
<td>0.66 (0.47, 0.92)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. MTX</td>
<td>0.88 (0.67, 1.15)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine</td>
<td>0.36 (0.18, 0.71)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine</td>
<td>0.35 (0.19, 0.63)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine</td>
<td>0.36 (0.19, 0.68)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine</td>
<td>0.48 (0.26, 0.89)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
Figure 27. Forest plot for network meta-analysis of discontinuations due to adverse events: TNF + MTX versus csDMARD

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adalimumab+MTX vs. MTX</td>
<td>0.97 (0.29, 3.22)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. MTX</td>
<td>1.08 (0.68, 1.73)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. MTX</td>
<td>0.81 (0.50, 1.29)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. MTX</td>
<td>3.03 (1.56, 5.90)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine</td>
<td>0.48 (0.12, 1.96)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine</td>
<td>0.53 (0.22, 1.29)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine</td>
<td>0.39 (0.16, 0.96)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine</td>
<td>1.48 (0.54, 4.06)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.

Certolizumab pegol. The C-OPERA trial (N=316) examined the combination of CZP plus MTX. At 2 years, the overall discontinuation rate for CZP plus MTX was 53.5 percent vs. 63.7 percent for MTX monotherapy (p=NR). Discontinuations attributable to adverse events and serious adverse events did not differ significantly between groups (low SOE). Similarly, the C-EARLY trial (N=879) observed a lower discontinuation rate for CZP plus MTX over 1 year (24.2% vs. 34.7%, p=NR) but no differences in discontinuations due to adverse events or differences in serious adverse events between groups. In NWMA, there was lower overall discontinuation for CZP plus MTX versus MTX monotherapy (RR, 0.64; 95% CI, 0.52 to 0.78) but no significant differences in discontinuations due to adverse events (Figure 26 and Figure 27, respectively).

Etanercept. Three trials compared ETN with MTX; one (N=542) compared combination therapy ETN plus MTX with MTX monotherapy; the two others (N=632 and N=26) compared ETN with MTX monotherapy. In the two larger trials, overall discontinuation rates were higher for the MTX-only group (12.7% vs. 10.2% and 40.5% vs. 25.6%); no significant differences in serious adverse events and discontinuations attributable to serious adverse events were observed in all three trials (low SOE). In NWMA, ETN plus MTX had a lower overall discontinuation rate than MTX monotherapy (RR, 0.66; 95% CI, 0.47 to 0.92) but no significant differences in discontinuation due to adverse events (Figure 26 and Figure 27, respectively).
Infliximab. Two trials assessed adverse events from combinations of IFX (3 mg/kg/8 weeks or 6 mg/kg/8 weeks) plus MTX compared with MTX monotherapy. The ASPIRE trial (N=1,049) found no significant differences in overall discontinuation rates (21.4% vs. 24.8% vs. 25.5%, p=NR), discontinuations attributable to adverse effects (9.5% vs. 9.6% vs. 3.2%, p=NR), and serious adverse events (11.0% vs. 14.0% vs. 14.0%, p=NR) (low SOE). Rates of serious infections, however, were higher in the IFX plus MTX groups than in the MTX monotherapy group (5.6%, 5.0%, 2.1%, p=0.02). Another smaller trial described lower overall discontinuation rates for IFX plus MTX than MTX monotherapy (6.7% vs. 14.3%, p=NR), one serious adverse event in the IFX plus MTX group (MTX related pneumonia), and similar side effects (benign infections, mild hepatotoxicity), but the sample was much smaller (N=44). In NWMA, there were no significant differences in overall discontinuation for IFX plus MTX, but there were higher discontinuations due to adverse events than MTX (RR, 3.03; 95% CI, 1.56 to 5.90) (Figure 26 and Figure 27, respectively).

TNF Biologic Versus csDMARD Combination Therapy

Adalimumab. The IMPROVED trial was a 2-year multicenter randomized single-blind trial (N=161) comparing ADA plus MTX with a combination of MTX, HCQ, and SSZ plus PRED in MTX nonresponders. Serious adverse events did not differ significantly (insufficient SOE). However, patients in the ADA plus MTX group experienced elevated liver enzymes at 4 percent and patients in the four-drug combination group at 8 percent (p=NR).

Infliximab. The SWEFOT trial was a multicenter randomized trial comparing MTX plus SSZ plus HCQ with IFX plus MTX over 1 year in MTX non responders. The IFX plus MTX group reported lower overall discontinuation than the csDMARD combination group (18.0% vs. 31.5%, p=0.014). Rates of serious adverse events (0.8% vs. 0.8%, p=NR) and discontinuation attributable to adverse events (7.8% vs. 10.8%, p=NR) were similar.

The NEO-RACo trial also found no significant differences in either discontinuation attributable to adverse events (2.0% vs. 0.0%, p=NR) or serious adverse events (6.0% vs. 8.0%, p=NR). Overall, the SOE was low for discontinuations due to adverse events and serious adverse events.

Single-Arm Studies: TNF Biologics only

A single-arm observational cohort study (N=12,656) in the Swedish Rheumatology Quality Register examined patients with incident RA and subsequent diagnosis of lymphoma. After adjustment for age, sex, and inflammatory activity during the first year of RA diagnosis, there was no increased lymphoma risk in patients who took a TNF inhibitor compared with those who did not take a TNF inhibitor (RR, 0.9; 95% CI, 0.9 to 1.9).

Non-TNF Biologics

Non-TNF Biologic Plus MTX Versus Either Non-TNF Biologic or MTX

Serious adverse events or discontinuations attributable to adverse events did not differ significantly between the non-TNF biologics in combination with MTX versus MTX monotherapy (low SOE for ABA, moderate SOE for RIT).

Abatacept. Two trials compared the combination of ABA plus MTX with MTX only: the AGREE trial (N=509) and the AVERT study (N=351). Both trials found no significant differences in overall discontinuation rates, discontinuation attributable to adverse events, or serious adverse events. In NWMA, the csDMARD ABA plus MTX had fewer overall
discontinuations than SSZ (RR 0.47; 95% CI, 0.24 to 0.92) and discontinuations due to adverse events (RR, 0.24; 95% CI, 0.09 to 0.61) (Figure 28 and Figure 29, respectively). There was no difference in overall discontinuation between ABA plus MTX and MTX alone, though ABA plus MTX had less discontinuation due to adverse events (RR, 0.49, 95% CI, 0.28 to 0.86).

**Figure 28. Forest plot for network meta-analysis of all discontinuations: Non-TNF + MTX versus csDMARD**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. MTX</td>
<td>0.85 (0.57, 1.28)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. MTX</td>
<td>1.02 (0.77, 1.35)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Sulfasalazine</td>
<td>0.47 (0.24, 0.92)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Sulfasalazine</td>
<td>0.56 (0.31, 1.02)</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
Rituximab. The 2-year IMAGE trial (N=755) randomized patients to RIT (1 g days 1 and 15) plus MTX (7.5 mg-20 mg/week) combination therapy, RIT (500 mg days 1 and 15) plus MTX (7.5 mg-20 mg/week) combination therapy, or MTX monotherapy. Overall discontinuation rates were 29 percent in the MTX monotherapy group compared with 15 percent in both RIT plus MTX combination therapy groups (p=NR). Discontinuation attributable to adverse events and serious adverse events did not differ across the groups.

Tocilizumab. Two RCTs, previously described in the csDMARDs versus non-TNF biologics section, examined discontinuation rates for patients receiving combination therapy with TCZ plus MTX and patients receiving either TCZ or MTX monotherapy: the FUNCTION 2-year trial (N=1,162) and the U-Act-Early 2-year trial (N=317). Overall discontinuation rates and discontinuation attributable to either adverse events (U-Act-Early: 8.5% vs. 9.7% vs. 7.4%, p=0.82) or serious adverse events (U-Act-Early: 16.0% vs. 18.4% vs. 12.0%, p=0.44) did not differ across these groups (moderate SOE).

Biologic Head to Head: TNF Versus Non-TNF
The ORBIT trial (N=329), an open-label noninferiority RCT, compared the non-TNF biologic RIT with the TNF, ADA or ETN, over 1 year rated high risk of bias. Overall discontinuations (18.8% vs. 17.7%, p=NR) and discontinuations attributable to adverse events (1.4% vs. 1.3%, p=NR) did not differ between the two groups. The RIT group, however, had higher rates of serious adverse events than the comparison group, primarily related to infections.
and neutropenia (25.7% vs. 17.2%, p=NR). The harms evidence was insufficient for head-to-head comparisons of TNF and non-TNF biologics.

In our NWMA of TNF versus non-TNF, ETN led to fewer overall discontinuations than ABA (RR, 0.38; 95% CI, 0.20 to 0.74) and discontinuations due to adverse events (RR, 0.35; 95% CI, 0.17 to 0.71) (Figure 30 and Figure 31, respectively). There were also higher rates of discontinuations due to adverse events with CZP plus MTX (RR, 2.21; 95% CI, 1.07 to 4.57) or IFX plus MTX (RR, 6.18; 95% CI, 2.59 to 14.72) than ABA plus MTX. ETN alone also had fewer overall discontinuations than TCZ (RR, 0.59; 95% CI, 0.35 to 0.98) and discontinuations due to adverse events (RR, 0.30; 95% CI, 0.14 to 0.63). There was less overall discontinuation for CZP plus MTX than TCZ plus MTX (RR 0.63, 95% CI, 0.44 to 0.90) and less discontinuation due to adverse events for ETN plus MTX than TCZ plus MTX (RR, 0.39; 95% CI, 0.20 to 0.74).

**Figure 30. Forest plot for network meta-analysis of all discontinuations: TNF versus non-TNF**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etanercept vs. Abatacept</td>
<td>0.38 (0.20, 0.74)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept+MTX</td>
<td>0.77 (0.43, 1.38)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept+MTX</td>
<td>0.75 (0.47, 1.18)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept+MTX</td>
<td>0.77 (0.45, 1.31)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept+MTX</td>
<td>1.03 (0.63, 1.69)</td>
</tr>
<tr>
<td>Etanercept vs. Tocilizumab</td>
<td>0.59 (0.35, 0.98)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Tocilizumab+MTX</td>
<td>0.64 (0.39, 1.07)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Tocilizumab+MTX</td>
<td>0.63 (0.44, 0.90)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Tocilizumab+MTX</td>
<td>0.64 (0.42, 1.00)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Tocilizumab+MTX</td>
<td>0.86 (0.58, 1.28)</td>
</tr>
</tbody>
</table>

More harms with second drug More harms with first drug

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
TNF Versus TNF

No direct evidence was available for TNF versus TNF. The SOE for all indirect estimates was low (downgrading for indirectness and imprecision in all cases). In NWMA, there were no differences detected in overall discontinuations. IFX plus MTX led to higher rates of overall discontinuations due to adverse events than both CZP plus MTX (RR, 2.80; 95% CI, 1.24 to 6.31) and ETN plus MTX (RR, 3.76; 95% CI, 1.66 to 8.51) (Figure 32 and Figure 33, respectively). Other comparisons shown below did not have significant differences.
Figure 32. Forest plot for network meta-analysis of all discontinuations: TNF versus TNF

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>0.97 (0.61, 1.54)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>1.00 (0.59, 1.70)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>1.34 (0.81, 2.20)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Certolizumab+MTX</td>
<td>1.03 (0.69, 1.52)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Certolizumab+MTX</td>
<td>1.38 (0.98, 1.93)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Etanercept+MTX</td>
<td>1.34 (0.87, 2.06)</td>
</tr>
</tbody>
</table>

More harms with second drug | More harms with first drug

95% CI = 95% confidence interval; MTX = methotrexate; RR = relative risk; TNF = tumor necrosis factor; vs. = versus.
Figure 33. Forest plot for network meta-analysis of discontinuations due to adverse events: TNF versus TNF

Non-TNF Versus Non-TNF

No direct evidence was available for non-TNF versus non-TNF. The SOE for all indirect estimates was low (downgrading for indirectness and imprecision in all cases). In NWMA, there were no differences detected in overall discontinuations between TCZ and ABA or TCZ and ABA with MTX (Figure 34 and Figure 35, respectively). Discontinuations due to adverse events were only higher for TCZ plus MTX than ABA plus MTX (RR, 4.25; 95% CI, 2.07 to 8.72) (Figure 35).
Combinations and Therapy Strategies

Long-term studies of combination strategies using multiple csDMARDs or csDMARD plus TNF biologics ultimately showed no differences in serious adverse events between immediate combination and step-up therapies (low SOE). The BeSt trial (N=508) examined four groups: (1)
sequential DMARD, starting with MTX; (2) stepped-up combination therapy with MTX followed by SSZ, HCQ, and prednisone; (3) initial combination therapy of MTX, SSZ, and tapered high-dose PRED; and (4) initial combination therapy with MTX and IFX. In general, discontinuation rates trended highest in group 2 (step-up combination therapy) after 5 years (12%, 22%, 15%, 9%, p=0.05). Serious adverse events did not differ significantly across the groups. At 10 years, there were also no significant adverse events across groups (events per 100 patient/years: 13.2, 10.9, 12.1, 13.4, p=0.47).

The GUEPARD study randomized MTX-naïve patients to 3 months of ADA plus MTX or MTX monotherapy. In patients who at 3 months did not respond to an initial strategy, investigators examined whether disease activity–driven treatment with TNF inhibitors was equally effective in controlling clinical symptoms and structural damage in both groups. Overall discontinuations trended higher for the ADA plus MTX initial strategy (15.2% vs. 9.4%, p=NR), but there were no significant differences in serious adverse events between groups. We rated this study high ROB after 12 weeks because of the risk of contamination bias given that patients could be switched to difference dosing and treatment regimens when low disease activity was achieved at 12 weeks and beyond.

The two year OPERA study of 180 early RA patients in Danish hospital-based clinics using a treat to target protocol found numerically lower overall discontinuations (10.1% vs 16.5%, p=NR) and lower serious adverse events (n= 4 vs. n=11, p=NR) in the ADA plus MTX strategy than the MTX plus placebo group.

The TEAR trial (N=755) randomized patients to four treatment arms: (1) immediate treatment with MTX plus ETN; (2) immediate treatment with MTX plus SSZ plus HCQ (triple therapy); (3) step-up from MTX to MTX plus ETN if DAS28-ESR was 3.2 or higher; and (4) step-up from MTX to triple therapy if DAS28-ESR was 3.2 or higher. We rated this trial as high ROB because overall discontinuation rates were high (up to 42%); the therapy groups did not differ, however, on this measure. In addition, adverse events did not differ significantly across the groups.

**KQ 4: Comparative Benefits and Harms in Subgroups of Patients**

For KQ 4, we were interested in differences in benefits and harms among subpopulations based on age, sex or gender, race or ethnicity, disease activity, prior therapies, concomitant therapies, and other serious medical conditions. For most of our eligible interventions and for most subgroups of interest, we did not find any comparative evidence. The available evidence was limited to post hoc subgroup analyses of some TNF biologics versus csDMARDs.

**Key Points**

- For most comparisons of interest, we did not find any eligible evidence on differences in benefits and harms among subpopulations.
- The available evidence is limited to post hoc analyses without statistical subgroup analyses. It provides no reliable information on differences among subpopulations.
- Evidence was insufficient to draw any conclusions about response rates between older and younger patients or about response rate and radiographic changes between people with different levels of disease activity who were taking MTX with or without a TNF biologic (ADA or IFX).
Evidence was insufficient to draw any conclusions about serious adverse events as defined by FDA between older and younger patients who were taking MTX or the TNF biologic ETN.

**Detailed Synthesis**

**Corticosteroids**
We found no eligible evidence on subgroups of interest.

**csDMARDs**
We found no eligible evidence on subgroups of interest.

**TNF Biologic Versus csDMARD Monotherapy**
Post hoc analyses of data from three RCTs provided information on some subgroups of interest. These analyses were limited to ADA plus MTX, ETN monotherapy, and IFX plus MTX compared with MTX monotherapy. Because of the post hoc nature of these analyses, results should be interpreted cautiously. None of these studies conducted subgroup analyses using tests of interaction.

*Adalimumab.* A post hoc subgroup analysis of the HOPEFUL 1 trial assessed the impact of patients’ disease activity on radiographic progression and remission. In multivariate regression analyses, low disease activity at baseline was statistically significantly associated with no radiographic progression (p=0.01) and with remission (p=0.02) in patients treated with MTX but not in those on ADA and MTX combination treatment (insufficient SOE). The analyses did not compare the two subgroups directly.

*Etanercept.* A descriptive, retrospective analysis of the ERA trial presented data on efficacy and serious adverse events in patients 65 years or older and those younger than 65 years of age. The investigators did not conduct any statistical subgroup analyses. After 24 months of ETN treatment, patients 65 years or older had lower ACR response rates than younger patients (ACR50, 22% vs. 54%; ACR70, 14% vs. 32%) (insufficient SOE). Likewise, older patients in the MTX group had lower ACR response rates than younger patients (ACR50, 31% vs. 43%; ACR70, 13% vs. 25%) (insufficient SOE). Older patients had substantially higher risks of serious adverse events than younger patients in the ETN group (32.1 events vs. 4.6 events per 100 patient-years) and in the MTX group (41.7 events vs. 7.2 events per 100 patient-years) (insufficient SOE). The specific serious adverse events were not described in the study.

*Infliximab.* A post hoc analyses of the ASPIRE trial found that progression of joint damage was related to patients’ disease activity in both the IFX plus MTX and the MTX monotherapy groups. Patients with low, moderate, and high disease activity, however, experienced less joint damage in the IFX plus MTX group than in the MTX monotherapy group (p=0.01) (insufficient SOE).

**Combinations and Therapy Strategies**
In post hoc subgroup analyses of the SWEFOT study, investigators determined the impact of obesity on treatment effects. The SWEFOT study compared triple therapy of synthetic DMARDs (MTX + SSZ + HZQ) with a combination therapy of IFX plus MTX. Post hoc subgroup analyses stratified patients into those with a body mass index (BMI) greater than 30, a BMI between 25 and 29.9, and those with normal weight and a BMI of less than 25. Among all
patients, normal-weight patients achieved higher rates of EULAR good-response at 24 months than obese patients (66% vs. 38%; OR, 3.2; 95% CI, 1.4 to 7.3). Likewise, normal-weight patients had higher rates of remission (52% vs. 15%; OR, 6.0; 95 CI%, 1.6 to 22.6) than obese patients. The study did not determine the effect of obesity on the comparative benefits and harms of these treatment regimens.
Discussion

Overview of Key Findings

We conducted a systematic review and network meta-analysis (NWMA) to update the 2012 review of the comparative effectiveness of drug therapies for rheumatoid arthritis (RA);\(^1\) in this report we focused solely on early RA in adults (within 1 year of diagnosis). The objective was to evaluate the comparative effectiveness and harms of monotherapies, combination therapies, and different treatment strategies. These therapies include several categories of drugs: (1) corticosteroids; (2) two classes of disease-modifying antirheumatic drugs (DMARDs)—conventional synthetic (cs) and targeted synthetic (ts) DMARDS; (3) two classes of biologic DMARDs—tumor necrosis factor (TNF) and non-TNF biologics; and (4) biosimilars. The drug classes and constituent drugs and their abbreviations/acronyms can be found in Table 1.

A total of 41 randomized controlled trials (RCTs) and 8 observational studies comprised the evidence base of this updated review. Table 12 summarizes our findings about benefits and harms and gives the strength of evidence grades (SOE, in bold) for three Key Questions (KQs) addressed by this report. Studies (n=2) or study outcomes rated high risk of bias were excluded from analyses and used only in sensitivity analyses for the network meta-analysis. Given that there were sparse data available about subgroups (KQ4), we present this information after the table. SOE grades reflect the level of certainty about conclusions drawn from findings; they are high, moderate, low, or insufficient. Detailed assessment of the SOE for KQ outcomes can be found in Appendix E.

Of specific interest are the following outcomes related to efficacy—disease activity, radiographic changes, functional capacity, and remission—and the following outcomes related to harms—overall discontinuations, discontinuations attributable to adverse effects and serious adverse events. The study population included patients with moderate to high disease activity.

### Table 12. Summary of findings about benefits and harms of treatments for early rheumatoid arthritis with strength of evidence grades

<table>
<thead>
<tr>
<th>Key Comparisons</th>
<th>Efficacy</th>
<th>Harms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strength of Evidence (in Bold)</td>
<td>Strength of Evidence (in Bold)</td>
</tr>
<tr>
<td><strong>Corticosteroids:</strong></td>
<td>Remission significantly higher in corticosteroid plus MTX combination therapy than MTX alone</td>
<td>No significant differences in serious adverse events</td>
</tr>
<tr>
<td></td>
<td><strong>Low:</strong> downgraded because open label design; high attrition; and not enough events to meet optimal information size</td>
<td><strong>Moderate:</strong> downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>Corticosteroid + csDMARD vs. csDMARDs</td>
<td>Disease activity and radiographic progression</td>
<td>No significant differences in discontinuation attributable to adverse effects</td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> both outcomes downgraded because open label design; high attrition; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
<td><strong>Low:</strong> downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> downgraded because open label design; high attrition; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
</tbody>
</table>

1. Reference not cited in the text.
<table>
<thead>
<tr>
<th>Key Comparisons</th>
<th>Efficacy</th>
<th>Harms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Corticosteroids:</strong></td>
<td><strong>ACR response, radiographic progression, or remission</strong></td>
<td>Discontinuation attributable to adverse effects</td>
</tr>
<tr>
<td>High-dose corticosteroid (≥250 mg) + MTX vs. IFX</td>
<td><strong>Insufficient:</strong> all outcomes downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
<td><strong>Insufficient:</strong> downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td><strong>Serious adverse events in methyl-PNL + MTX vs. IFX</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> downgraded because open label design, and not enough events to meet optimal information size</td>
<td><strong>Insufficient:</strong> downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td><strong>Corticosteroids:</strong></td>
<td><strong>ACR response, remission, or functional capacity</strong></td>
<td>Discontinuation attributable to adverse effects</td>
</tr>
<tr>
<td>High-dose corticosteroid (≥250 mg) + MTX vs. MTX</td>
<td><strong>Insufficient:</strong> downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
<td><strong>Insufficient:</strong> downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
<td><strong>Serious adverse events in methyl-PNL + MTX vs. MTX</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
<td><strong>Insufficient:</strong> downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td><strong>csDMARDs:</strong></td>
<td><strong>Disease activity in PNL + SSZ vs. PNL + MTX</strong></td>
<td>Discontinuation attributable to adverse effects</td>
</tr>
<tr>
<td>csDMARDs vs. csDMARDs</td>
<td><strong>Insufficient (based on RCTs):</strong> downgraded because high attrition; large baseline differences between groups; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td><strong>Insufficient:</strong> downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient (based on observational evidence):</strong> downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td><strong>Insufficient (based on observational evidence):</strong> downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td><strong>Remission in PNL + SSZ vs. PNL + MTX</strong></td>
<td><strong>Remission in PNL + SSZ vs. PNL + MTX</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient:</strong> downgraded because high attrition; direction of effect varies; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td><strong>Insufficient:</strong> downgraded because high attrition; direction of effect varies; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy</td>
<td>Harms</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>Strength of Evidence</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>Functional capacity in PNL + SSZ vs. PNL + MTX</td>
<td><strong>Insufficient</strong>: downgraded because high attrition; large baseline differences between groups; and not enough events to meet optimal information size</td>
<td></td>
</tr>
<tr>
<td>Functional capacity in SSZ vs. MTX</td>
<td><strong>Insufficient (based on observational evidence)</strong>: downgraded because high risk of confounding by indication</td>
<td></td>
</tr>
<tr>
<td>csDMARDs:</td>
<td>No significant differences in response or remission in MTX + SSZ vs. MTX</td>
<td>No significant differences in discontinuation attributable to adverse effects in MTX + SSZ vs. MTX</td>
</tr>
<tr>
<td><strong>csDMARD Combination Therapy vs. csDMARD Monotherapy</strong></td>
<td><strong>Low (based on RCTs)</strong>: downgraded because open label design; high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td><strong>Low (based on RCTs)</strong>: Downgraded because open label design; high attrition; and imprecision</td>
</tr>
<tr>
<td>Insufficient (based on observational evidence):</td>
<td>Downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td>Insufficient (based on observational evidence): Downgraded because high risk of selection bias for treatment discontinuation and confounding by indication; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>No significant differences in functional capacity for MTX + SSZ vs. MTX at 1 year or 5 years, or for comparisons of PNL + MTX + SSZ + HCQ vs. MTX or SSZ</td>
<td><strong>Low</strong>: downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
<td>No significant differences in serious adverse events in MTX + SSZ vs. MTX</td>
</tr>
<tr>
<td>Radiographic progression for csDMARD combination therapy vs. csDMARD monotherapy</td>
<td>Insufficient: downgraded because open label design; high attrition; and large CIs cross appreciable benefits or harms</td>
<td><strong>Low</strong>: Downgraded because open label design, and high attrition</td>
</tr>
<tr>
<td>csDMARDs:</td>
<td>ACR response and remission significantly higher, radiographic progression less, and functional capacity significantly improved with ADA + MTX vs. ADA or with ADA vs. MTX</td>
<td>No significant differences in discontinuation because adverse events or serious adverse events for ADA + MTX vs. ADA or for ADA vs. MTX</td>
</tr>
<tr>
<td><strong>csDMARDs vs. TNF Biologics</strong></td>
<td><strong>Moderate</strong>: downgraded because high attrition</td>
<td><strong>Moderate</strong>: downgraded because high attrition</td>
</tr>
<tr>
<td>ADA + MTX vs. ADA or ADA vs. MTX</td>
<td>No significant differences in ACR response or remission for ADA + MTX vs. ABA or for ABA vs. MTX</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events for ADA + MTX vs. ABA or for ABA vs. MTX</td>
</tr>
<tr>
<td>csDMARDs:</td>
<td>Low: both outcomes downgraded because high attrition</td>
<td>Low: both outcomes downgraded because high attrition</td>
</tr>
<tr>
<td><strong>csDMARDs vs. Non-TNF Biologics</strong></td>
<td>No significant differences in functional capacity for ABA + MTX vs. ABA or for ABA vs. MTX</td>
<td></td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy Strength of Evidence</td>
<td>Harms Strength of Evidence</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>csDMARDs:</td>
<td>Remission significantly higher for TCZ + MTX vs. TCZ and TCZ vs. MTX</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events for TCZ + MTX vs. TCZ or for TCZ vs. MTX</td>
</tr>
<tr>
<td>csDMARDs vs. Non-TNF Biologics</td>
<td>Low: downgraded because large CIs cross appreciable benefits or harms</td>
<td>Moderate: both outcomes downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ or TCZ vs. MTX</td>
<td>Functional capacity for TCZ + MTX vs. TCZ and TCZ vs. MTX</td>
<td>Discontinuation attributable to adverse effects or serious adverse events for TOF + MTX vs. MTX or TOF</td>
</tr>
<tr>
<td></td>
<td>Insufficient: downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
<td>Insufficient: both outcomes downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td>Disease activity for TCZ + MTX vs. TCZ and TCZ vs. MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insufficient: downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>csDMARDs:</td>
<td>ACR response, disease activity, remission, and radiographic progression for TOF + MTX vs. MTX or TOF</td>
<td></td>
</tr>
<tr>
<td>csDMARD vs. tsDMARD</td>
<td>Insufficient: all outcomes downgraded because high attrition; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional capacity for TOF + MTX vs. MTX or TOF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insufficient: downgraded because large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>Biologics</td>
<td>Functional capacity significantly improved for ADA + MTX vs. MTX</td>
<td>No significant differences in discontinuation because adverse events for ADA + MTX vs. MTX</td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs. csDMARD Monotherapy</td>
<td>Moderate: downgraded because high attrition</td>
<td>Low: both outcomes downgraded because high attrition; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td>ADA + MTX vs. MTX</td>
<td>ACR response significantly higher with ADA + MTX vs. MTX</td>
<td>No significant differences in serious adverse events for ADA + MTX vs. MTX</td>
</tr>
<tr>
<td></td>
<td>Low: downgraded because high attrition, and large CIs cross appreciable benefits or harms</td>
<td>Low: both outcomes downgraded because high attrition; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td></td>
<td>Remission significantly higher with ADA + MTX vs. MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low: both outcomes downgraded because high attrition, and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Radiographic progression less with ADA + MTX vs. MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low: downgraded because high attrition, and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy</td>
<td>Harms</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td><strong>Biologics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs.</td>
<td>ACR response significantly</td>
<td>No significant differences</td>
</tr>
<tr>
<td>csDMARD Monotherapy</td>
<td>higher and radiographic</td>
<td>in discontinuation</td>
</tr>
<tr>
<td></td>
<td>progression less for CZP + MTX</td>
<td>because adverse</td>
</tr>
<tr>
<td></td>
<td>Low: both outcomes</td>
<td>effects or serious</td>
</tr>
<tr>
<td></td>
<td>downgraded because high</td>
<td>adverse events</td>
</tr>
<tr>
<td></td>
<td>attrition; large CIs; and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>not enough events to meet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>optimal information size</td>
<td></td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>Remission significantly</td>
<td></td>
</tr>
<tr>
<td></td>
<td>higher and functional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>capacity improved for CZP +</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTX Low: both outcomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>downgraded because high</td>
<td></td>
</tr>
<tr>
<td></td>
<td>attrition; large CIs; and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>not enough events to meet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>optimal information size</td>
<td></td>
</tr>
<tr>
<td><strong>Biologics</strong></td>
<td>ACR response significantly</td>
<td>No significant differences</td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs.</td>
<td>higher and radiographic</td>
<td>in discontinuation</td>
</tr>
<tr>
<td>csDMARD Monotherapy</td>
<td>progression less for ETN + MTX</td>
<td>because adverse</td>
</tr>
<tr>
<td></td>
<td>Moderate: both outcomes</td>
<td>effects or serious</td>
</tr>
<tr>
<td></td>
<td>downgraded because medium</td>
<td>adverse events</td>
</tr>
<tr>
<td></td>
<td>level of study limitations</td>
<td></td>
</tr>
<tr>
<td>ETN + MTX or ETN vs. MTX</td>
<td>Remission rates significantly</td>
<td>Low: both outcomes</td>
</tr>
<tr>
<td></td>
<td>higher for ETN + MTX and ETN</td>
<td>downgraded because medium</td>
</tr>
<tr>
<td></td>
<td>vs. MTX Low: downgraded</td>
<td>level of study limitations,</td>
</tr>
<tr>
<td></td>
<td>because medium level of study</td>
<td>and not enough events to</td>
</tr>
<tr>
<td></td>
<td>limitations, and not enough</td>
<td>meet optimal information</td>
</tr>
<tr>
<td></td>
<td>events to meet optimal</td>
<td>size</td>
</tr>
<tr>
<td></td>
<td>information size</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional capacity mixed</td>
<td>Low: downgraded because</td>
</tr>
<tr>
<td></td>
<td>for ETN + MTX and ETN vs. MTX</td>
<td>direction of effect varies,</td>
</tr>
<tr>
<td></td>
<td>Low: downgraded because</td>
<td>and large CIs</td>
</tr>
<tr>
<td></td>
<td>medium level of study</td>
<td></td>
</tr>
<tr>
<td></td>
<td>limitations, and not enough</td>
<td></td>
</tr>
<tr>
<td></td>
<td>events to meet optimal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>information size</td>
<td></td>
</tr>
<tr>
<td><strong>Biologics</strong></td>
<td>Remission rates significantly</td>
<td>No significant differences</td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs.</td>
<td>higher and functional</td>
<td>in discontinuation</td>
</tr>
<tr>
<td>csDMARD Monotherapy</td>
<td>capacity greater for IFX + MTX</td>
<td>because adverse</td>
</tr>
<tr>
<td></td>
<td>Low: both outcomes</td>
<td>effects or serious</td>
</tr>
<tr>
<td></td>
<td>downgraded because medium</td>
<td>adverse events</td>
</tr>
<tr>
<td></td>
<td>level of study limitations</td>
<td></td>
</tr>
<tr>
<td>IFX + MTX vs. MTX</td>
<td>Disease activity and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>radiographic progression for</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFX + MTX vs. MTX Insufficient: both outcomes downgraded because not enough events to meet optimal information size; direction of effect varies; and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td><strong>Biologics</strong></td>
<td>Disease activity, radiographic</td>
<td>Serious adverse events</td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs.</td>
<td>progression, or remission</td>
<td>Insufficient: downgraded</td>
</tr>
<tr>
<td>csDMARD Combination Therapy</td>
<td>for ADA + MTX vs. MTX + PRED +</td>
<td>because high attrition;</td>
</tr>
<tr>
<td>(e.g., triple therapy)</td>
<td>HCQ + SSZ Insufficient: all outcomes downgraded because high attrition; not enough events to meet optimal information size; and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>ADA + MTX vs. MTX + PRED + HCQ +</td>
<td>Functional capacity for ADA +</td>
<td></td>
</tr>
<tr>
<td>SSZ Insufficient:</td>
<td>MTX vs. MTX + PRED + HCQ + SSZ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional capacity for ADA +</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTX vs. MTX + PRED + HCQ + SSZ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insufficient: downgraded</td>
<td></td>
</tr>
<tr>
<td></td>
<td>because high attrition, and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>not enough events to meet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>optimal information size</td>
<td></td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy Strength of Evidence</td>
<td>Harms Strength of Evidence</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs. csDMARD Combination Therapy (e.g., triple therapy)</td>
<td>ACR response significantly higher for IFX + MTX vs. MTX + SSZ + HCQ <strong>Low</strong>: downgraded because medium level of study limitations</td>
<td>No significant differences in discontinuation attributable to either adverse effects or serious adverse events <strong>Low</strong>: both outcomes downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>IFX + MTX vs. MTX + SSZ + HCQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF Biologics: TNF Biologic vs. csDMARD Combination Therapy (triple therapy)</td>
<td>No significant differences in ACR response, radiographic progression, or remission for IFX + MTX + SSZ + HCQ + PRED vs. MTX + SSZ + HCQ + PRED <strong>Low</strong>: all outcomes downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events <strong>Low</strong>: both outcomes downgraded because large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td>IFX + MTX + SSZ + HCQ + PRED vs. MTX + SSZ + HCQ + PRED</td>
<td>No significant differences in functional capacity for IFX + MTX + SSZ + HCQ + PRED vs. MTX + SSZ + HCQ + PRED <strong>Low</strong>: downgraded because not enough events to meet optimal information size</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events</td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-TNF Biologics: Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>Disease activity significantly improved and remission rates higher for ABA + MTX vs. MTX <strong>Moderate</strong>: both outcomes downgraded because high attrition, and large baseline differences between groups</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events <strong>Low</strong>: both outcomes downgraded because high attrition</td>
</tr>
<tr>
<td>ABA + MTX vs. MTX</td>
<td>Radiographic progression significantly less for ABA + MTX vs. MTX</td>
<td>Functional capacity mixed for ABA + MTX vs. MTX <strong>Low</strong>: downgraded because high attrition; direction of effect varies; large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy Strength of Evidence</td>
<td>Harms Strength of Evidence</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-TNF Biologics: Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>Disease activity significantly improved and radiographic progression less for RIT + MTX vs. MTX</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events</td>
</tr>
<tr>
<td></td>
<td><strong>Moderate</strong>: both outcomes downgraded because not enough events to meet optimal information size</td>
<td><strong>Moderate</strong>: both outcomes downgraded because not enough events to meet optimal information size</td>
</tr>
<tr>
<td>RIT + MTX vs. MTX</td>
<td>Remission rates significantly higher for RIT + MTX vs. MTX</td>
<td><strong>Moderate</strong>: both outcomes downgraded because not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td><strong>Moderate</strong>: downgraded because single-study body of evidence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional capacity significantly improved for RIT + MTX vs. MTX</td>
<td></td>
</tr>
<tr>
<td>Biologics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-TNF Biologics: Non-TNF Biologic vs. csDMARD Monotherapy</td>
<td>Radiographic progression less for TCZ + MTX vs. MTX</td>
<td>No significant differences in discontinuation attributable to adverse effects or serious adverse events</td>
</tr>
<tr>
<td></td>
<td><strong>Moderate</strong>: downgraded because large baseline differences between groups</td>
<td><strong>Moderate</strong>: both outcomes downgraded because medium level of study limitations</td>
</tr>
<tr>
<td>TCZ + MTX vs. MTX</td>
<td>Remission significantly higher for TCZ + MTX vs. MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Low</strong>: downgraded because medium level of study limitations, and large confidence intervals cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disease activity and functional capacity for TCZ + MTX vs. MTX</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient</strong>: both outcomes downgraded because direction of effect varies, and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>Biologics: TNF vs. Non-TNF Biologics</td>
<td>Functional capacity significantly improved for RIT vs. ADA or ETN</td>
<td>Discontinuation attributable to adverse effects or serious adverse events</td>
</tr>
<tr>
<td></td>
<td><strong>Low</strong>: downgraded because no ITT analysis, and high risk of selection bias for treatment discontinuation and confounding by indication</td>
<td><strong>Insufficient</strong>: both outcomes downgraded because no ITT analysis; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
</tr>
<tr>
<td></td>
<td>Disease activity or remission for RIT vs. ADA or ETN</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Insufficient</strong>: both outcomes downgraded because no ITT analysis; large CIs cross appreciable benefits or harms; and not enough events to meet optimal information size</td>
<td></td>
</tr>
<tr>
<td>Key Comparisons</td>
<td>Efficacy Strength of Evidence</td>
<td>Harms Strength of Evidence</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Combination Strategies: 1: Sequential monotherapy starting with MTX vs. 2: Step-up combination therapy vs. 3: Combination with high-dose tapered PRED vs. 4: Combination therapy with IFX</td>
<td>Disease activity significantly more improved for strategy 3 (combination therapy with high dose tapered PRED) and strategy 4 (combination therapy with IFX) than with either strategy 1 (sequential monotherapy) or 2 (step-up therapy) in short term (1 year), but no significant differences in long term (4 or 10 years) <strong>Moderate</strong>: downgraded because large CIs cross appreciable benefits or harms</td>
<td>No significant differences in serious adverse events <strong>Low</strong>: downgraded because large CIs cross appreciable benefits or harms</td>
</tr>
<tr>
<td></td>
<td>No significant differences in long term radiographic progression (10 years) <strong>Moderate</strong>: downgraded because large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No significant differences in long term remission (4 or 10 years) <strong>Moderate</strong>: downgraded because large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No significant differences in long term functional capacity (2, 5, or 10 years) <strong>Low</strong>: downgraded because not enough events to meet optimal information size, and large CIs cross appreciable benefits or harms</td>
<td></td>
</tr>
<tr>
<td>Combination Strategies: 1: Immediate MTX + ETN vs. 2: Immediate MTX + SSZ + HCQ vs. 3: Step-up MTX to combo MTX + ETN vs. 4: Step-up MTX to combo MTX + SSZ + HCQ</td>
<td>Disease activity, remission, radiographic progression, or functional capacity for immediate combination therapy (MTX + ETN) vs. step-up triple therapy (MTX + SSZ + HCQ) <strong>Insufficient</strong>: all outcomes downgraded because high attrition; no ITT analysis; and large CIs cross appreciable benefits or harms</td>
<td>Discontinuation attributable to adverse effects or serious adverse events <strong>Insufficient</strong>: both outcomes downgraded because high attrition; no ITT analysis; and large CIs cross appreciable benefits or harms</td>
</tr>
</tbody>
</table>
## Key Comparisons

<table>
<thead>
<tr>
<th>Efficacy</th>
<th>Harms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combination Strategies:</strong></td>
<td></td>
</tr>
<tr>
<td>ADA + MTX adjusted based on DAS vs. MTX</td>
<td>Disease activity, remission, or radiographic progression for ADA + MTX adjusted based on DAS vs. MTX</td>
</tr>
<tr>
<td><strong>Insufficient:</strong> all outcomes downgraded because high attrition, and large CIs cross appreciable benefits or harms</td>
<td>Discontinuation attributable to adverse effects or serious adverse events</td>
</tr>
<tr>
<td>Functional capacity for ADA + MTX adjusted based on DAS vs. MTX</td>
<td>Insufficient: both outcomes downgraded because high attrition, and large CIs cross appreciable benefits or harms</td>
</tr>
</tbody>
</table>

**Strength of Evidence**

- ADA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; CI = confidence interval; csDMARD = conventional synthetic DMARD; CZP = certolizumab pegol; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; HCQ = hydroxychloroquine; IFX = infliximab; ITT = intent-to-treat; MTX = methotrexate; obs = observational; PRED = prednisone; RIT = rituximab; SSZ = sulfasalazine; TCZ = tocilizumab; TNF = tumor necrosis factor; TOF = tofacitininb; tsDMARD = targeted synthetic DMARD; vs. = versus.

Existing comparative evidence for our review was diverse. It included comparisons of monotherapies, combination therapies, triple therapy (methotrexate [MTX], sulfasalazine [SSZ], hydroxychloroquine [HCQ]), and treatment strategies. Additionally, the drug classes spanned corticosteroids, csDMARDs, tsDMARDs, TNF biologic DMARDs, and non-TNF biologic DMARDs. No studies on the use of biosimilar DMARD agents in early RA were included in this report because they did not fit the inclusion criteria.

For corticosteroids and csDMARDs, the evidence allowed us to draw some conclusions for early RA. Corticosteroids, in combination with MTX, led to higher remission rates than MTX alone for MTX naïve patients with moderate to severe disease; results were mixed, however, for radiographic progression, health-related quality of life (HRQOL), and functional capacity. There were no significant differences in serious adverse events and discontinuations attributable to adverse events between these two treatment regimens. The corticosteroids used were heterogeneous and included varying doses of prednisone (PRED), prednisolone, and methylprednisolone regimens.

Studies of csDMARD therapies mainly examined SSZ and MTX. Comparisons of combination therapy with monotherapy found no differences in disease activity, functional capacity, serious adverse events, or discontinuations attributable to adverse events.

Although several biologic agents are available, the head-to-head evidence remains limited. Moderate strength of evidence supports combination therapy of adalimumab (ADA) plus MTX versus ADA only for several outcomes; specifically, ADA plus MTX led to higher American College of Rheumatology (ACR) response rates, higher remission rates, and less radiographic progression than ADA monotherapy. There were no significant differences in serious adverse events or discontinuations attributable to adverse events between these two medication regimens. Our NWMA also found significantly higher ACR50 response rates and less radiographic progression following use of ADA plus MTX versus ADA monotherapy. The data showed that both TNF biologics (ADA, etanercept [ETN], or infliximab [IFX]), but not non-TNF biologics (abatacept [ABA] or tocilizumab [TCZ]), in combination with MTX have higher ACR50 treatment response than biologic monotherapy. The results of comparative NWMA for overall
The evidence comparing TNF biologics (ADA, certolizumab pegol [CZP], ETN, or IFX) plus MTX with MTX monotherapy generally showed higher remission rates, better functional capacity, and less radiographic progression for the combination medications. Serious adverse events or discontinuations did not differ significantly. Similar findings were also noted for non-TNF biologics (ABA, rituximab [RIT], or TCZ) in combination with MTX. Head-to-head evidence for biologics is limited to one trial, which found no significant differences in disease activity and remission compared with TNF biologics (ADA or ETN).

Combination therapies with csDMARD triple therapy (MTX plus SSZ plus HCQ) compared with TNF biologics (either ADA or IFX) plus MTX found no significant differences in remission or radiographic changes. Rates of adverse events did not differ. In terms of treatment strategies, the BeSt study assessed several treatment strategies for early RA; the investigators included sequential monotherapy, step-up combination therapy, combination therapy with tapered PRED, and combination therapy with IFX. Over the long term (i.e., 10 years), radiographic progression, remission, and functional capacity did not differ across the arms of the trial.

Subpopulation data were limited to post hoc analyses. For most comparisons, we did not find eligible evidence on the benefits and harms among subpopulations.

Findings in Relationship to What Is Already Known

We conducted a systematic review and NWMA to update the 2012 review of the comparative effectiveness of drug therapies for RA; in this report we focused solely on early RA in adults (within 1 year of diagnosis). All of the early RA studies included patients with moderate to high disease activity. In a clinical setting, patients with early RA may present with varying levels of severity. Also, the studies did not consistently parse out which patients had tried one or more therapies and which ones were treatment naïve.

Our results go further than treatment recommendations for early RA from the ACR and the European League against Rheumatism (EULAR) and support additional therapies for patients with moderate to high levels of disease activity. The ACR and EULAR task force both support a treat-to-target approach over a nontargeted approach with the goal of achieving remission or low disease activity. The BeSt and FIN-RACo trials used a treat-to-target approach, and their results support the ACR and EULAR recommendations in this respect.

The ACR guidelines recommend csDMARD monotherapy (MTX preferred) instead of double or triple csDMARD therapy in patients who have never taken a csDMARD. If disease activity remains moderate or high, despite csDMARD monotherapy, then the ACR recommends double or triple csDMARD therapy or a TNF or non-TNF biologic DMARD (with or without MTX). Our evidence was insufficient to support one DMARD over another (e.g., csDMARDs, biologic DMARDs). However, we found that when biologics were used in combination with MTX therapy, patients achieved lower disease activity, higher functional capacity, and higher remission rates than with monotherapy alone. The difference between the results of our findings and the ACR guidelines may be due to a few reasons. First, all of our studies included patients with moderate to high disease activity at baseline. Patients with early RA in a clinical setting may present with less disease severity and prior history with MTX could vary. Additionally, this report assessed comparative effectiveness based on current available evidence and included secondary longer time points when available. Clinical practice guidelines use systematic reviews as evidence and if evidence is not enough they may consider other resources. The ACR based
their recommendations on a consideration of the balance of relative benefits and harms of the 
treatment options as well as expert opinion and preferences.

Although the evidence for the effectiveness of MTX plus biologics in early RA is favorable, 
it is not the standard of care for a number of reasons. First, some data indicate that certain 
patients will do well on MTX monotherapy, but no information is available about how to identify 
or predict these patients. Second, many insurers require MTX failure as a prerequisite to add a 
bioptic (probably based on the effectiveness of MTX). Third, patients may be wary of a 
combination therapy approach in early disease (e.g., cost, side effects, injections). Additionally, 
patients must balance the potentially higher efficacy of multiple drugs with the burden and 
potential for increased risk.

The EULAR task force advocates starting with csDMARDs as first-line therapy in the 
absence of poor prognostic factors (e.g., high disease activity, early joint damage, autoantibody 
positivity) in early RA.43 When poor prognostic factors are present, the task force advocates for 
adding a TNF or non-TNF biologic to a csDMARD. This guideline group regards all biologic 
DMARDs as similarly effective and safe after csDMARD monotherapy failure. Our findings 
harmonize with EULAR’s guidelines recommending combination therapy with a biologic as 
first-line therapy for patients with poor prognostic factors. The evidence we found comparing 
combinations of biologics and MTX with either biologic or MTX monotherapies (N=10 studies) 
in patients with early RA and poor prognostic factors reported that patients receiving 
combination therapies achieved higher remission rates.12-15, 17, 32-34, 37, 41 However, we had no 
available studies that specifically examined the effect of therapies in patients with early RA and 
less severe disease activity to patients with early RA plus poor prognostic factors.

Applicability

Although we derived our evidence primarily from RCTs that typically enrolled a discrete 
population and were conducted under ideal situations, the findings from observational and 
efficacy trials were generally consistent. However, the observational and noncontrolled studies 
reported higher discontinuation rates. For example, one observational study of MTX versus 
MTX plus SSZ in a SSZ-resistant population had overall discontinuation rates ranging from 33.9 
percent to 50.0 percent at 1 year due to either side effects or lack of response.26 A second 
observational study of MTX versus SSZ reported similar reasons for discontinuation.28 
Discontinuation rates from clinical trials were generally lower than 20 percent. The higher 
discontinuation rates in observational studies may reflect real-world settings as compared with 
the tighter adherence in a controlled clinical trial. The observational studies in this report 
describing harms were rated as medium to high risk of bias. Higher quality observational studies 
may affect the estimates of these results.

The range of mean (or median) disease durations across all 49 included studies was 2 weeks 
to 12 months. All our included studies enrolled patients with moderate to high disease activity at 
baseline as measured with mean or median Disease Activity Scale (DAS) 28 scores, ranging 
broadly from 3.4 to 7.1 (DAS ranges from 0 to 10; 3.2 is a threshold for low disease activity; 
more than 5.1 is considered high disease activity). More than one-half of the patient population 
were women; the mean age range was 46 to 64 years. Study durations ranged from 6 months to 
15 years.

In addition, trials comparing corticosteroids used varying doses and tapering strategies. 
Similarly, MTX dosing ranged from 7.5 mg per week to 25 mg per week. This degree of
heterogeneity did not allow for suitable evidence comparison, but it may be typical of common clinical practices.

As stated previously, subpopulation studies of differences in benefits and harms were mostly lacking. The data were sparse for any comparative differences in serious infections and malignancies in this early RA population. The evidence was limited to post hoc subgroup analyses from studies comparing TNF biologics with csDMARDs.

**Contextual Questions**

During the review process, we flagged studies for their relevance to the contextual questions during the review process and we also supplemented this evidence base with a targeted literature search.

**Contextual Question 1: Does treatment of early RA improve disease trajectory and disease outcomes compared with the trajectory or outcomes of treatment of established RA?**

Structural damage occurs early in active RA, and early DMARD treatment improves the long-term outcome of the disease.² In prospective studies of early RA, approximately 75 percent of patients have joint erosions or develop initial erosions within the first 2 years of symptom onset.¹⁶⁹ In a review of five delayed treatment trials, RA patients treated immediately at presentation had improved patient function and reduced radiographic progression than patients whose treatment was delayed.⁴⁴ For the majority of these trials, the average disease duration at initial presentation was 12 months or less. Few other data support these results, however, because it is now thought to be unethical to withhold treatment from patients in early active RA.

The ultimate treatment goal for RA is sustained remission. However, less than 50 percent of all RA patients who achieve remission remain in remission 1 year later.¹⁷⁰ Achieving remission earlier in the disease trajectory is important to achieving goals such as reduction of joint damage and disability.¹⁷¹ In one observational study of 871 women with RA, patients who achieved remission less than 5 years after diagnosis were able to maintain remission, while patients who first achieved remission 5 or more years after diagnosis were not able to do so.¹⁷² A meta-analysis of data on RA patients from 14 RCTs identified that one strong predictor of a beneficial response to therapy was a shorter disease duration at treatment initiation.¹⁷³

**Contextual Question 2: What barriers prevent individuals with early RA from obtaining access to indicated drug therapies?**

One qualitative research study of health care stakeholders, including general practitioners, rheumatologists, hospital representatives, and members of a rheumatology society (N=34), identified key barriers to accessing appropriate (or any) care for early RA. Important barriers included lack of access to primary health care services because of travel distance, difficulties of making an RA diagnosis in primary care, difficulties in accessing biologics and obtaining insurer approval of biologics, and lack of access to specialty care, especially in rural areas.⁴⁵

A cross-sectional study of 4,037 RA patients identified clinical situations in which rheumatologists elected to continue monitoring RA in patients with moderate or high disease activity rather than adjusting their DMARD therapy.¹⁷⁴ Several circumstances prompted this practice: patient preference not to adjust therapy, insufficient time to assess response to recently
initiated DMARD treatment, noninflammatory musculoskeletal pain contributing to a high DAS28 score, costs, and reimbursement issues.

Another qualitative study of rheumatologists and nurses (N=32) explored barriers hindering the use of intensive combination treatment strategies in early RA patients. Several important barriers were identified: contraindications (e.g., patients with coexisting conditions, older patients), increased risk of side effects and related complications, and patients’ resistance to therapies.46

Patients face high out-of-pocket expenses for RA therapies. In a retrospective analysis of the Medical Expenditure Panel Survey, mean out-of-pocket expenses were $274.99 per monthly prescription.47 This figure was lower for privately insured and publicly insured patients than for those who were uninsured. Higher out-of-pocket expenses were found among patients who were uninsured, female, and diagnosed with other conditions in addition to RA.

In a 12-month observational study using Marketscan Research databases (N=26,911), the research team examined risk factors for noninitiation of DMARDs in patients with newly diagnosed RA. 175 Early RA patients were followed for 12 months after diagnosis. More than one-third of patients did not start DMARD therapy within that first year. After multivariate adjustment, risk factors for DMARD noninitiation included older age (85 years or older); high Deyo-Charlson Comorbidity Index score; and the presence of gastrointestinal disorders, cardiac conditions, hypertension, osteoarthritis, or respiratory infections.

**Limitations**

Our review update has some limitations. No consensus exists on the definition of early RA. Moreover, criteria used in the literature for defining populations with early RA are variable. A recent task force of RA experts recommended defining early RA as no more than 1 year of diagnosed disease duration.43 For this review, we defined populations with early RA as having a diagnosed disease duration limited to 1 year or less and included mixed population studies if >50 percent of the study populations had an early RA diagnosis. It is possible that patients described in this way may have longer disease (symptoms).

Additional evidence on treatment comparisons might be gained by expanding the definition to 2 years. However, requiring a diagnosed disease duration of 1 year or less is in line with current clinical practice. In reviewing our literature, we identified but excluded 7 studies (reported in 10 articles) of adults with a duration of RA between 1 and 2 years from diagnosis. On brief review of the 7 studies, findings did not differ from the current report.

For several of the studies evaluating corticosteroids, drug dosing was heterogeneous. This factor limited our ability to draw conclusions from comparisons of these agents. Similarly, in csDMARD comparisons, MTX dosing varied from 7.5 mg to 25 mg weekly.

Few data were available about subgroups that are of interest to this field; typically, we found only limited data on age. Evidence was limited for the tsDMARD class and nonexistent for biosimilars in the early RA population. Although existing evidence of biologics in combination with MTX shows that this regimen can improve disease activity, we do not know whether starting treatment with a biologic rather than a csDMARD improves long-term prognosis of RA.

Because of a lack of head-to-head trials, we often had to rely on results from the NWMAs to estimate the comparative effectiveness of interventions of interest for treating patients with early RA. Network (sometimes referred to as indirect or mixed) meta-analyses are an important analytic tool in the absence of direct head-to-head evidence, but they also have limitations. The “transitivity assumption” relies on the premise that any patient in the network would be equally
likely to have received any of the treatments in the network. It is difficult to assess this assumption when no direct head-to-head studies are available and estimates are based exclusively on indirect comparisons. In the case of our NWMAs, most comparisons were based on a “star network” with MTX as the common comparator. A star network indicates a dearth of head-to-head studies directly comparing interventions. Most effect estimates, therefore, were derived from indirect comparisons rather than mixed treatment comparisons. Although we carefully assessed the clinical heterogeneity of all trials included in the network meta-analyses to ensure that they were as homogenous as possible, we were not able to statistically assess the assumption of homogeneity (and transitivity) for most comparisons. Furthermore, NWMAs often yield estimates with wide confidence intervals that encompass clinically relevant benefits or harms for both drugs (or combinational therapies) that are being compared. Such inconclusive results should not be misinterpreted as evidence for no difference in benefits or harms. In general, these limitations are reflected in the strength of evidence ratings.

**Research Needs**

Future research should help clinicians and researchers draw stronger conclusions on the comparative effectiveness and harms of medications for patients with early RA. Multiple established therapies exist for early RA, but comparative evidence is badly needed. Studies comparing therapy options in patients diagnosed with early RA who have different degrees of disease activity or poor prognostic factors would be helpful in the clinical setting.

Also, at least some, or perhaps many therapies for early RA may be initially effective, but longer-term effects have not been well studied. Studies with longer treatment periods and followup of 5 years or longer would provide better information on adherence and adverse events. Registry data have the potential to include real-world populations with data on long-term effects and follow-up. They would also yield insights as to whether starting with a biologic improves long-term prognosis of RA.

Most studies that we used for this review evaluated csDMARD and biologic medications. FDA has approved several biosimilars, but because they have not been studied specifically among early RA patients, we could not include any studies of them in this review. Per FDA guidance, efficacy outcomes for the biosimilars are based on extrapolation from studies in several indications and may not be specifically studied in RA, either early or late. Four of 39 studies reporting radiographic outcomes described MRI findings. This is an evolving technology.

Analyses of subpopulations based on age and coexisting medical conditions (hepatitis C, congestive heart failure, diabetes, and cancer) would also be helpful for clinicians and patients newly diagnosed with RA. Currently, treatment selection based on benefits and harms is difficult in these populations. Additionally, patient-centered research is needed with appropriate use of patient-reported outcomes and other patient-generated health data so that results are truly reflective of patient preferences and desires.

**Conclusions**

For patients with early RA, qualitative and network meta-analyses suggest that the combination of MTX with TNF or non-TNF biologics improves disease activity and remission when compared with monotherapy with a biologic or csDMARD. This comprehensive review found similar adverse events and discontinuation rates for csDMARDs, TNF biologics, and non-TNF biologics in studies ranging in length from 6 months to 15 years.
## Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation or Acronym</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>abatacept</td>
</tr>
<tr>
<td>ACR</td>
<td>American College of Rheumatology (20/50/70 = 20%/50%/70% improvement)</td>
</tr>
<tr>
<td>ADA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>adalimumab</td>
</tr>
<tr>
<td>AE(s)</td>
<td>adverse event(s) (S = serious)</td>
</tr>
<tr>
<td>AGREE</td>
<td>Abatacept trial to Gauge Remission and joint damage progression in methotrexate-naive patients with Early Erosive rheumatoid arthritis</td>
</tr>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>ALT</td>
<td>alanine transaminase</td>
</tr>
<tr>
<td>ASPIRE</td>
<td>Active-controlled Study of Patients receiving Infliximab for the treatment of Rheumatoid arthritis of Early onset trial</td>
</tr>
<tr>
<td>AST</td>
<td>aspartate aminotransferase</td>
</tr>
<tr>
<td>AVERT</td>
<td>Assessing Very Early Rheumatoid arthritis Treatment trial</td>
</tr>
<tr>
<td>Avg</td>
<td>average</td>
</tr>
<tr>
<td>BARFOT</td>
<td>Better Anti Rheumatic FarmacOTherapy trial</td>
</tr>
<tr>
<td>BeSti</td>
<td>Dutch acronym for Behandel-Strategieen, “treatment strategies” trial</td>
</tr>
<tr>
<td>biwkly</td>
<td>biweekly</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>BRAF-Mdq</td>
<td>Bristol Rheumatoid Arthritis Fatigue – Multidimensional Questionnaire</td>
</tr>
<tr>
<td>C-OPERA</td>
<td>Certolizumab-Optimal Prevention of joint damage for Early RA trial</td>
</tr>
<tr>
<td>CAMERA-II</td>
<td>Computer Assisted Management in Early Rheumatoid Arthritis trial-II</td>
</tr>
<tr>
<td>CARDERA</td>
<td>Combination Anti-Rheumatic Drugs in Early Rheumatoid Arthritis trial</td>
</tr>
<tr>
<td>CareRA</td>
<td>Care for Early RA trial</td>
</tr>
<tr>
<td>CCP</td>
<td>cyclic citrullinated peptide</td>
</tr>
<tr>
<td>COBRA</td>
<td>Dutch acronym for Combinatietherapie Bij Reumatoide Arthritis trial</td>
</tr>
<tr>
<td>combo</td>
<td>combination therapy</td>
</tr>
<tr>
<td>COMET</td>
<td>Combination of Methotrexate and Etanercept in Active Early Rheumatoid Arthritis trial</td>
</tr>
<tr>
<td>CQ</td>
<td>Contextual Question</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>CS</td>
<td>corticosteroid(s)</td>
</tr>
<tr>
<td>csDMARD</td>
<td>conventional synthetic DMARD</td>
</tr>
<tr>
<td>CZP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>certolizumab pegol</td>
</tr>
<tr>
<td>D</td>
<td>day(s)</td>
</tr>
<tr>
<td>d/c</td>
<td>discontinuation</td>
</tr>
<tr>
<td>DAS</td>
<td>Disease Activity Score (based on 44 joints)</td>
</tr>
<tr>
<td>DAS28</td>
<td>Disease Activity Score in 28 joints (**-ESR = using ESR; **+-CRP = using CRP)</td>
</tr>
<tr>
<td>DAS28-4 ESR</td>
<td>Disease Activity Score in 28 joints with 4 variables including ESR</td>
</tr>
<tr>
<td>DMARD</td>
<td>disease modifying antirheumatic drug</td>
</tr>
<tr>
<td>Enbrel ERA</td>
<td>Enbrel Early RA trial</td>
</tr>
<tr>
<td>EPC</td>
<td>Evidence-based Practice Center</td>
</tr>
<tr>
<td>ESR</td>
<td>erythrocyte sedimentation rate</td>
</tr>
<tr>
<td>ETN&lt;sup&gt;b&lt;/sup&gt;</td>
<td>etanercept</td>
</tr>
<tr>
<td>EULAR</td>
<td>European League against Rheumatism</td>
</tr>
<tr>
<td>EQ-5D or EuroQoL</td>
<td>European Quality of Life Questionnaire-5 Dimensions</td>
</tr>
<tr>
<td>EuroQoL VAS</td>
<td>European Quality of Life Visual Analogue Scale</td>
</tr>
<tr>
<td>FDA</td>
<td>U.S. Food and Drug Administration</td>
</tr>
<tr>
<td>FIN-RACo</td>
<td>Finnish Rheumatoid Arthritis Combination Therapy trial</td>
</tr>
<tr>
<td>FUNCTION</td>
<td>trial whose acronym not described</td>
</tr>
<tr>
<td>G1, 2…</td>
<td>group 1, 2…</td>
</tr>
<tr>
<td>g/day</td>
<td>grams per day</td>
</tr>
<tr>
<td>GC</td>
<td>glucocorticoid</td>
</tr>
<tr>
<td>GI</td>
<td>gastrointestinal</td>
</tr>
<tr>
<td>GOL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>golimumab</td>
</tr>
<tr>
<td>GUEPARD</td>
<td>French acronym for Guerir la PolyArthrite Rhumatoide Debutante [cure early RA] trial</td>
</tr>
<tr>
<td>Abbreviation or Acronym</td>
<td>Meaning</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>HAQ</td>
<td>Health Assessment Questionnaire</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>Disability Index of the Health Assessment Questionnaire (HAQ)</td>
</tr>
<tr>
<td>HCO²</td>
<td>hydroxychloroquine</td>
</tr>
<tr>
<td>HIT HARD</td>
<td>High Induction Therapy with Anti-Rheumatic Drugs trial</td>
</tr>
<tr>
<td>HOPEFUL 1</td>
<td>Human anti-TNF monoclonal antibody Outcome study for the Persistent EFficacy Under allocation to treatment strategies in early RA</td>
</tr>
<tr>
<td>HR</td>
<td>hazard ratio</td>
</tr>
<tr>
<td>HRQOL</td>
<td>health related quality of life</td>
</tr>
<tr>
<td>i.m.</td>
<td>intramuscular</td>
</tr>
<tr>
<td>IDEA</td>
<td>Infliximab as Induction Therapy in Early Rheumatoid Arthritis trial</td>
</tr>
<tr>
<td>IFX²</td>
<td>infliximab</td>
</tr>
<tr>
<td>IMAGE</td>
<td>International study in Methotrexate-Naive Patients Investigating Rituximab’s Efficacy</td>
</tr>
<tr>
<td>IMPROVED</td>
<td>Induction therapy with Methotrexate and Prednisone in Rheumatoid or Very Early arthritic Disease trial</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>ITT</td>
<td>intention to treat</td>
</tr>
<tr>
<td>IV</td>
<td>intravenous</td>
</tr>
<tr>
<td>Kg</td>
<td>kilograms</td>
</tr>
<tr>
<td>KQ</td>
<td>Key Question</td>
</tr>
<tr>
<td>LEF³</td>
<td>leflunomide</td>
</tr>
<tr>
<td>Max</td>
<td>maximum</td>
</tr>
<tr>
<td>Methyl-PNL d</td>
<td>methylprednisolone</td>
</tr>
<tr>
<td>mg</td>
<td>milligrams</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>mos</td>
<td>months</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>mTSS</td>
<td>Modified Sharp/van der Heijde Method for Scoring Radiographs score</td>
</tr>
<tr>
<td>MTX c</td>
<td>methotrexate</td>
</tr>
<tr>
<td>N</td>
<td>number</td>
</tr>
<tr>
<td>NA</td>
<td>not applicable</td>
</tr>
<tr>
<td>NEO-RACo</td>
<td>New Finnish RA Combination Therapy trial</td>
</tr>
<tr>
<td>NOR-DMARD</td>
<td>Norwegian Antirheumatic Drug Register</td>
</tr>
<tr>
<td>NR</td>
<td>not reported</td>
</tr>
<tr>
<td>nRCT</td>
<td>nonrandomized controlled trial</td>
</tr>
<tr>
<td>NS</td>
<td>not significant</td>
</tr>
<tr>
<td>NWMA</td>
<td>network meta-analysis(es)</td>
</tr>
<tr>
<td>OPERA</td>
<td>OPTimized treatment algorithm in Early RA</td>
</tr>
<tr>
<td>OPTIMA</td>
<td>Optimal Protocol for Treatment Initiation with Methotrexate and Adalimumab trial</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>ORBIT</td>
<td>Optimal Management of patients with rheumatoid arthritis who Require Biologic Therapy trial</td>
</tr>
<tr>
<td>PCORI</td>
<td>Patient-Centered Outcomes Research Institute</td>
</tr>
<tr>
<td>PICOTS</td>
<td>population, intervention/exposure, comparator, outcomes, time frames, country and clinical settings, and study design</td>
</tr>
<tr>
<td>PNL d</td>
<td>prednisolone</td>
</tr>
<tr>
<td>PRED d</td>
<td>prednisone</td>
</tr>
<tr>
<td>PREMIER</td>
<td>trial whose acronym not described</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</td>
</tr>
<tr>
<td>PRO</td>
<td>patient-reported outcome</td>
</tr>
<tr>
<td>PROSPERO</td>
<td>International Prospective Register of Systematic Reviews</td>
</tr>
<tr>
<td>PROWD</td>
<td>PRevention of Work Disability trial</td>
</tr>
<tr>
<td>pt-years</td>
<td>patient-years</td>
</tr>
<tr>
<td>QOL</td>
<td>quality of life</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>RA</td>
<td>rheumatoid arthritis</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>Abbreviation or Acronym</td>
<td>Meaning</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>RF</td>
<td>rheumatoid factor</td>
</tr>
<tr>
<td>RIT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>rituximab</td>
</tr>
<tr>
<td>ROB</td>
<td>Risk of bias</td>
</tr>
<tr>
<td>RR</td>
<td>risk ratio</td>
</tr>
<tr>
<td>SAR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>sarilumab</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>SF-12</td>
<td>Short Form Survey (MCS = Mental Component Score; PCS = Physical Component Score)</td>
</tr>
<tr>
<td>SF-36</td>
<td>Short Form 36 Health Survey (MCS = Mental Component Score; PCS = Physical Component Score)</td>
</tr>
<tr>
<td>SHS</td>
<td>Sharp/van der Heijde Method for Scoring Radiographs</td>
</tr>
<tr>
<td>SIR</td>
<td>standardized incidence ratio</td>
</tr>
<tr>
<td>SMD</td>
<td>standardized mean difference</td>
</tr>
<tr>
<td>SOE</td>
<td>strength of evidence</td>
</tr>
<tr>
<td>SR</td>
<td>systematic review</td>
</tr>
<tr>
<td>SSZ&lt;sup&gt;c&lt;/sup&gt;</td>
<td>sulfasalazine</td>
</tr>
<tr>
<td>SWEFOT</td>
<td>Swedish Pharmacotherapy Study</td>
</tr>
<tr>
<td>TB</td>
<td>tuberculosis</td>
</tr>
<tr>
<td>TCZ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>tocilizumab</td>
</tr>
<tr>
<td>TEAR</td>
<td>Treatment of Early Aggressive Rheumatoid Arthritis Trial</td>
</tr>
<tr>
<td>TNF</td>
<td>tumor necrosis factor</td>
</tr>
<tr>
<td>TOF&lt;sup&gt;e&lt;/sup&gt;</td>
<td>tofacitinib</td>
</tr>
<tr>
<td>tREACH</td>
<td>treatment in the Rotterdam Early Arthritis Cohort trial</td>
</tr>
<tr>
<td>tsDMARD</td>
<td>targeted synthetic DMARD</td>
</tr>
<tr>
<td>TSU</td>
<td>tight step-up</td>
</tr>
<tr>
<td>U-Act-Early</td>
<td>trial whose acronym not described</td>
</tr>
<tr>
<td>ULN</td>
<td>upper limit of normal</td>
</tr>
<tr>
<td>URTI</td>
<td>upper respiratory tract infection</td>
</tr>
<tr>
<td>UTI</td>
<td>urinary tract infection</td>
</tr>
<tr>
<td>VAS</td>
<td>visual analogue scale</td>
</tr>
<tr>
<td>vs.</td>
<td>versus</td>
</tr>
<tr>
<td>WPS-RA</td>
<td>Work Productivity Survey - Rheumatoid Arthritis</td>
</tr>
<tr>
<td>yrs</td>
<td>years</td>
</tr>
<tr>
<td>95% CI</td>
<td>95% confidence interval</td>
</tr>
</tbody>
</table>

<sup>a</sup> Non-TNF Biologics  
<sup>b</sup> TNF Biologics  
<sup>c</sup> csDMARDs  
<sup>d</sup> Corticosteroids  
<sup>e</sup> tsDMARDs
Appendix A. Search Strings

PubMed: April 11, 2017 Original Searches

Results: 1778 imported after removing duplicates

The original search retrieved an original total of 1934 records, and after initial removal of 149 duplicates, 1785 records were left. Further deduplication of seven records left 1778 records for literature review.

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Search &quot;arthritis, rheumatoid&quot;[MeSH Terms] OR &quot;rheumatoid arthritis&quot;[All Fields]</td>
<td>130439</td>
</tr>
<tr>
<td>#2</td>
<td>Search (&quot;Adrenal Cortex Hormones&quot;[MeSH] OR corticosteroid&quot;)</td>
<td>312282</td>
</tr>
<tr>
<td>#3</td>
<td>Search (Methylprednisolone OR prednisone OR prednisolone)</td>
<td>106322</td>
</tr>
<tr>
<td>#4</td>
<td>Search (Hydroxychloroquine OR Leflunomide OR Methotrexate OR Sulfasalazine)</td>
<td>56751</td>
</tr>
<tr>
<td>#5</td>
<td>Search (Adalimumab OR &quot;certolizumab pegol&quot; OR etanercept OR golimumab OR infliximab OR Abatacept OR tocilizumab OR rituximab OR Tofacitinib OR Sarilumab OR Baricitinib OR Sirukumab)</td>
<td>39136</td>
</tr>
<tr>
<td>#6</td>
<td>Search (amjevita OR Inflectra OR Erelzi)</td>
<td>32</td>
</tr>
<tr>
<td>#7</td>
<td>Search (#2 OR #3 OR #4 OR #5 OR #6)</td>
<td>464392</td>
</tr>
<tr>
<td>#8</td>
<td>Search (#1 AND #7)</td>
<td>23197</td>
</tr>
<tr>
<td>#9</td>
<td>Search (letter[pt] OR newspaper article[pt] OR editorial[pt] OR comment[pt])</td>
<td>1557410</td>
</tr>
<tr>
<td>#10</td>
<td>Search (#8 NOT #9)</td>
<td>21103</td>
</tr>
<tr>
<td>#11</td>
<td>Search (#8 NOT #9) Filters: English</td>
<td>16989</td>
</tr>
<tr>
<td>#12</td>
<td>Search (#8 NOT #9) Filters: Humans; English</td>
<td>14387</td>
</tr>
<tr>
<td>#13</td>
<td>Search (#8 NOT #9) Filters: Humans; English; Adult: 19+ years</td>
<td>9107</td>
</tr>
<tr>
<td>#15</td>
<td>Search (&quot;review&quot;[Publication Type] AND &quot;systematic&quot;[tiab]) OR &quot;systematic review&quot;[All Fields] OR (&quot;review literature as topic&quot;[MeSH] AND &quot;systematic&quot;[tiab]) OR &quot;meta-analysis&quot;[Publication Type] OR &quot;meta-analysis as topic&quot;[MeSH Terms] OR &quot;meta-analysis&quot;[All Fields]</td>
<td>201594</td>
</tr>
<tr>
<td>#17</td>
<td>Search (13 AND 14)</td>
<td>1303</td>
</tr>
<tr>
<td>#18</td>
<td>Search (13 AND 15)</td>
<td>106</td>
</tr>
<tr>
<td>#19</td>
<td>Search (13 AND 16)</td>
<td>2867</td>
</tr>
<tr>
<td>#20</td>
<td>Search (13 AND 14) Filters: Publication date from 2010/07/01</td>
<td>529</td>
</tr>
<tr>
<td>#21</td>
<td>Search (13 AND 15) Filters: Publication date from 2010/07/01</td>
<td>65</td>
</tr>
<tr>
<td>#22</td>
<td>Search (13 AND 16) Filters: Publication date from 2010/07/01</td>
<td>1340</td>
</tr>
<tr>
<td>#23</td>
<td>Search (#20 OR #21 OR #22) Filters: Publication date from 2010/07/01</td>
<td>1785</td>
</tr>
</tbody>
</table>
PubMed: October 5, 2017 Update Searches

Results: 124 imported after removing duplicates

The update search retrieved an original total of 205 records, and after the removal of 81 duplicates, 124 records were left for literature review.

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Search &quot;arthritis, rheumatoid&quot;[MeSH Terms] OR &quot;rheumatoid arthritis&quot;[All Fields]</td>
<td>132907</td>
</tr>
<tr>
<td>#2</td>
<td>Search (&quot;Adrenal Cortex Hormones&quot;[MeSH] OR corticosteroid*)</td>
<td>317011</td>
</tr>
<tr>
<td>#3</td>
<td>Search (Methylprednisolone OR prednisone OR prednisolone)</td>
<td>107804</td>
</tr>
<tr>
<td>#4</td>
<td>Search (Hydroxychloroquine OR Leflunomide OR Methotrexate OR Sulfasalazine)</td>
<td>57816</td>
</tr>
<tr>
<td>#5</td>
<td>Search ((Adalimumab OR &quot;certolizumab pegol&quot; OR etanercept OR golimumab OR infliximab OR Abatacept OR tocilizumab OR rituximab OR Tofacitinib OR Sarilumab))</td>
<td>40963</td>
</tr>
<tr>
<td>#6</td>
<td>Search ((amjevita OR Inflectra OR Erelzi))</td>
<td>37</td>
</tr>
<tr>
<td>#7</td>
<td>Search (#2 OR #3 OR #4 OR #5 OR #6)</td>
<td>472306</td>
</tr>
<tr>
<td>#8</td>
<td>Search (#1 AND #7)</td>
<td>23796</td>
</tr>
<tr>
<td>#9</td>
<td>Search (letter[pt] OR newspaper article[pt] OR editorial[pt] OR comment[pt])</td>
<td>1589852</td>
</tr>
<tr>
<td>#10</td>
<td>Search (#8 NOT #9)</td>
<td>21674</td>
</tr>
<tr>
<td>#11</td>
<td>Search (#8 NOT #9) Filters: English</td>
<td>17534</td>
</tr>
<tr>
<td>#12</td>
<td>Search (#8 NOT #9) Filters: Humans; English</td>
<td>14911</td>
</tr>
<tr>
<td>#13</td>
<td>Search (#8 NOT #9) Filters: Humans; English; Adult: 19+ years</td>
<td>9490</td>
</tr>
<tr>
<td>#15</td>
<td>Search (&quot;review&quot;[Publication Type] AND &quot;systematic&quot;[tiab]) OR &quot;systematic review&quot;[All Fields] OR (review literature as topic)[MeSH] AND &quot;systematic&quot;[tiab] OR &quot;meta-analysis&quot;[Publication Type] OR &quot;meta-analysis as topic&quot;[MeSH Terms] OR &quot;meta-analysis&quot;[All Fields])</td>
<td>215879</td>
</tr>
<tr>
<td>#17</td>
<td>Search (#13 AND #14)</td>
<td>1393</td>
</tr>
<tr>
<td>#18</td>
<td>Search (#13 AND #15)</td>
<td>116</td>
</tr>
<tr>
<td>#19</td>
<td>Search (#13 AND #16)</td>
<td>3040</td>
</tr>
<tr>
<td>#20</td>
<td>Search (#13 AND #14) Filters: Publication date from 2016/10/01 to 2017/12/31</td>
<td>74</td>
</tr>
<tr>
<td>#21</td>
<td>Search (#13 AND #15) Filters: Publication date from 2016/10/01 to 2017/12/31</td>
<td>10</td>
</tr>
<tr>
<td>#22</td>
<td>Search (#13 AND #16) Filters: Publication date from 2016/10/01 to 2017/12/31</td>
<td>138 (121 imported)</td>
</tr>
</tbody>
</table>
Embase: April 11-12, 2017 Original Searches

Results: 1171 imported after removing duplicates

The original search retrieved an original total of 1413 records, and after initial removal of 101 duplicates, 1312 records were left. Further removal of another 141 duplicates left 1171 records for literature review.

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>'rheumatoid arthritis'/exp OR 'rheumatoid arthritis'</td>
<td>196,665</td>
</tr>
<tr>
<td>#2</td>
<td>'adrenal cortex hormones' OR corticosteroid*</td>
<td>282,225</td>
</tr>
<tr>
<td>#3</td>
<td>methylprednisolone OR prednisone OR prednisolone</td>
<td>320,723</td>
</tr>
<tr>
<td>#4</td>
<td>hydroxychloroquine OR leflunomide OR methotrexate OR sulfasalazine</td>
<td>173,827</td>
</tr>
<tr>
<td>#5</td>
<td>adalimumab OR 'certolizumab pegol' OR etanercept OR golimumab OR infliximab OR abatacept OR tocilizumab OR rituximab OR tofacitinib OR sarilumab OR Baricitinib OR sirukumab</td>
<td>115,006</td>
</tr>
<tr>
<td>#6</td>
<td>amjevita OR inflectra OR erelzi</td>
<td>168</td>
</tr>
<tr>
<td>#7</td>
<td>#2 OR #3 OR #4 OR #5 OR #6</td>
<td>703,322</td>
</tr>
<tr>
<td>#8</td>
<td>#1 AND #7</td>
<td>56,311</td>
</tr>
<tr>
<td>#9</td>
<td>#8 AND ('editorial'/it OR 'letter'/it)</td>
<td>4,240</td>
</tr>
<tr>
<td>#10</td>
<td>#8 NOT #9</td>
<td>52,071</td>
</tr>
<tr>
<td>#11</td>
<td>#10 AND ((adult]/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim OR [young adult]/lim)</td>
<td>17,805</td>
</tr>
<tr>
<td>#12</td>
<td>#10 AND ((adult]/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim OR [young adult]/lim) AND [humans]/lim AND [english]/lim</td>
<td>15,753</td>
</tr>
<tr>
<td>#13</td>
<td>'randomized controlled trial'/exp OR 'single blind procedure'/exp OR 'double blind procedure'/exp OR 'random allocation'/exp</td>
<td>540,522</td>
</tr>
<tr>
<td>#14</td>
<td>'systematic review'/exp OR 'systematic review (topic)'/exp OR 'meta-analysis (topic)'/exp OR 'meta analysis (topic)'/exp OR 'meta analysis'/exp</td>
<td>236,374</td>
</tr>
<tr>
<td>#15</td>
<td>'case control study'/exp OR 'cohort analysis'/exp OR 'epidemiological study' OR 'cross-sectional study'/exp OR 'organizational case study' OR 'crossover procedure'/exp OR 'seroepidemiologic study' OR 'epidemiology'/exp OR 'multicenter study'/exp OR 'multicenter study (topic)'/exp OR 'evaluation research'/exp</td>
<td>3,034,615</td>
</tr>
<tr>
<td>#16</td>
<td>#12 AND #13</td>
<td>1,426</td>
</tr>
<tr>
<td>#17</td>
<td>#12 AND #14</td>
<td>139</td>
</tr>
<tr>
<td>#18</td>
<td>#12 AND #15</td>
<td>4,073</td>
</tr>
<tr>
<td>#19</td>
<td>#16 AND [2010-2017]/py</td>
<td>692</td>
</tr>
<tr>
<td>#20</td>
<td>#17 AND [2010-2017]/py</td>
<td>113</td>
</tr>
<tr>
<td>#21</td>
<td>#18 AND [2010-2017]/py</td>
<td>2,799</td>
</tr>
<tr>
<td>#22</td>
<td>#19 AND [medline]/lim</td>
<td>456</td>
</tr>
<tr>
<td>#23</td>
<td>#20 AND [medline]/lim</td>
<td>45</td>
</tr>
<tr>
<td>#24</td>
<td>#21 AND [medline]/lim</td>
<td>1,659</td>
</tr>
<tr>
<td>#25</td>
<td>#19 NOT #22</td>
<td>236</td>
</tr>
<tr>
<td>#26</td>
<td>#20 NOT #23</td>
<td>68</td>
</tr>
<tr>
<td>#27</td>
<td>#21 NOT #24</td>
<td>1,140</td>
</tr>
<tr>
<td>#28</td>
<td>#25 OR #26 OR #27</td>
<td>1,312</td>
</tr>
</tbody>
</table>

A-3
**Embase: October 5, 2017 Update Searches**

**Results: 280 imported after removing duplicates**

The update search retrieved an original total of 356 records, and after initial removal of 21 duplicates, 335 records were left. Further removal of another 55 duplicates left 280 records for literature review.

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>'rheumatoid arthritis'/exp OR 'rheumatoid arthritis'</td>
<td>201,832</td>
</tr>
<tr>
<td>#2</td>
<td>'adrenal cortex hormones' OR corticosteroid*</td>
<td>290,037</td>
</tr>
<tr>
<td>#3</td>
<td>methylprednisolone OR prednisone OR prednisolone</td>
<td>329,261</td>
</tr>
<tr>
<td>#4</td>
<td>hydroxychloroquine OR leflunomide OR methotrexate OR sulfasalazine</td>
<td>178,657</td>
</tr>
<tr>
<td>#5</td>
<td>adalimumab OR 'certolizumab pegol' OR etanercept OR golimumab OR infliximab OR abatacept OR tocilizumab OR rituximab OR tofacitinib OR sarilumab</td>
<td>121,123</td>
</tr>
<tr>
<td>#6</td>
<td>amjevita OR inflectra OR erekz</td>
<td>215</td>
</tr>
<tr>
<td>#7</td>
<td>#2 OR #3 OR #4 OR #5 OR #6</td>
<td>724,071</td>
</tr>
<tr>
<td>#8</td>
<td>#1 AND #7</td>
<td>58,100</td>
</tr>
<tr>
<td>#9</td>
<td>#8 AND ('editorial'/it OR 'letter'/it)</td>
<td>4,339</td>
</tr>
<tr>
<td>#10</td>
<td>#8 NOT #9</td>
<td>53,761</td>
</tr>
<tr>
<td>#11</td>
<td>#10 AND [(adult)/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim OR [young adult]/lim]</td>
<td>18,752</td>
</tr>
<tr>
<td>#12</td>
<td>#10 AND [(adult)/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim OR [young adult]/lim] AND [humans]/lim AND [english]/lim</td>
<td>16,684</td>
</tr>
<tr>
<td>#13</td>
<td>'randomized controlled trial'/exp OR 'single blind procedure'/exp OR 'double blind procedure'/exp OR 'random allocation'/exp</td>
<td>571,238</td>
</tr>
<tr>
<td>#14</td>
<td>'systematic review'/exp OR 'systematic review (topic)'/exp OR 'meta analysis (topic)'/exp OR 'meta analysis'/exp</td>
<td>257,410</td>
</tr>
<tr>
<td>#15</td>
<td>'case control study'/exp OR 'cohort analysis'/exp OR 'epidemiological study' OR 'cross-sectional study'/exp OR 'organizational case study' OR 'crossover procedure'/exp OR 'seroepidemiologic study' OR 'epidemiology'/exp OR 'multicenter study'/exp OR 'multicenter study (topic)'/exp OR 'evaluation research'/exp</td>
<td>3,192,637</td>
</tr>
<tr>
<td>#16</td>
<td>#12 AND #13</td>
<td>1,501</td>
</tr>
<tr>
<td>#17</td>
<td>#12 AND #14</td>
<td>149</td>
</tr>
<tr>
<td>#18</td>
<td>#12 AND #15</td>
<td>4,459</td>
</tr>
<tr>
<td>#19</td>
<td>#16 AND 2017:py</td>
<td>94</td>
</tr>
<tr>
<td>#20</td>
<td>#17 AND 2017:py</td>
<td>11</td>
</tr>
<tr>
<td>#21</td>
<td>#18 AND 2017:py</td>
<td>464</td>
</tr>
<tr>
<td>#22</td>
<td>#19 AND [medline]/lim</td>
<td>38</td>
</tr>
<tr>
<td>#23</td>
<td>#20 AND [medline]/lim</td>
<td>0</td>
</tr>
<tr>
<td>#24</td>
<td>#21 AND [medline]/lim</td>
<td>148</td>
</tr>
<tr>
<td>#25</td>
<td>#19 NOT #22</td>
<td>56 (27 imported)</td>
</tr>
<tr>
<td>#26</td>
<td>#20 NOT #23</td>
<td>11 (11 imported)</td>
</tr>
<tr>
<td>#27</td>
<td>#21 NOT #24</td>
<td>316 (297 imported)</td>
</tr>
</tbody>
</table>
Cochrane Library: April 12, 2017 Original Searches

Results: 563 imported after removing duplicates

The original search retrieved an original total of 1067 records, and after initial removal of 3 duplicates, 1064 records were left. Further removal of another 501 duplicates left 563 records for literature review.

<table>
<thead>
<tr>
<th>ID</th>
<th>Search</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>[mh &quot;arthritis, rheumatoid&quot;] or &quot;rheumatoid arthritis&quot;</td>
<td>9745</td>
</tr>
<tr>
<td>#2</td>
<td>[mh &quot;Adrenal Cortex Hormones&quot;] or corticosteroid</td>
<td>20270</td>
</tr>
<tr>
<td>#3</td>
<td>Methylprednisolone or prednisone or prednisolone</td>
<td>14571</td>
</tr>
<tr>
<td>#4</td>
<td>Hydroxychloroquine or Leflunomide or Methotrexate or Sulfasalazine</td>
<td>8664</td>
</tr>
<tr>
<td>#5</td>
<td>Adalimumab or &quot;certolizumab pegol&quot; or etanercept or golimumab or infliximab or Abatacept or tocilizumab or rituximab or Tofacitinib or Sanilumab or Baricitinib or Sirukumab</td>
<td>6633</td>
</tr>
<tr>
<td>#6</td>
<td>amjevita or Inflectra or Erelzi</td>
<td>8</td>
</tr>
<tr>
<td>#7</td>
<td>#2 or #3 or #4 or #5 or #6</td>
<td>41459</td>
</tr>
<tr>
<td>#8</td>
<td>#1 and #7</td>
<td>4125</td>
</tr>
<tr>
<td>#9</td>
<td>&quot;randomized controlled trial&quot;:pt or &quot;randomized controlled trial as topic&quot;:pt or &quot;single-blind method&quot;:pt or &quot;double-blind method&quot;:pt or &quot;random allocation&quot;:pt</td>
<td>420356</td>
</tr>
<tr>
<td>#10</td>
<td>(review and systematic) or &quot;systematic review&quot; or (&quot;review literature as topic&quot; and systematic) or &quot;meta-analysis&quot;</td>
<td>63119</td>
</tr>
<tr>
<td>#11</td>
<td>[mh &quot;Cohort Studies&quot;] or [mh &quot;Epidemiologic Studies&quot;] or [mh &quot;Follow-up Studies&quot;] or &quot;prospective cohort&quot; or [mh &quot;prospective studies&quot;] or (prospective* and cohort and (study or studies))</td>
<td>148804</td>
</tr>
<tr>
<td>#12</td>
<td>#8 and #9 Publication Year from 2010 to 2017</td>
<td>700</td>
</tr>
<tr>
<td>#13</td>
<td>#8 and #10 Publication Year from 2010 to 2017</td>
<td>320</td>
</tr>
<tr>
<td>#14</td>
<td>#8 and #11 Publication Year from 2010 to 2017</td>
<td>261</td>
</tr>
<tr>
<td>#15</td>
<td>#12 or #13 or #14</td>
<td>1067</td>
</tr>
</tbody>
</table>

(1064 imported)
Cochrane Library: October 5, 2017 Update Searches

Results: 21 imported after removing duplicates

The update search retrieved an original total of 79 records, and after initial removal of one duplicate, 78 records were left. Further removal of another 57 duplicates left 21 records for literature review.

<table>
<thead>
<tr>
<th>ID</th>
<th>Search</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>[mh &quot;arthritis, rheumatoid&quot;] or &quot;rheumatoid arthritis&quot;</td>
<td>10184</td>
</tr>
<tr>
<td>#2</td>
<td>[mh &quot;Adrenal Cortex Hormones&quot;] or corticosteroid</td>
<td>21528</td>
</tr>
<tr>
<td>#3</td>
<td>Methylprednisolone or prednisone or prednisolone</td>
<td>15642</td>
</tr>
<tr>
<td>#4</td>
<td>Hydroxychloroquine or Leflunomide or Methotrexate or Sulfasalazine</td>
<td>9289</td>
</tr>
<tr>
<td>#5</td>
<td>Adalimumab or &quot;certolizumab pegol&quot; or etanercept or golimumab or infliximab or Abatacept or tocilizumab or rituximab or Tofacitinib or Sarilumab</td>
<td>7486</td>
</tr>
<tr>
<td>#6</td>
<td>amjevita or Inflectra or Erelzi</td>
<td>11</td>
</tr>
<tr>
<td>#7</td>
<td>#2 or #3 or #4 or #5 or #6</td>
<td>44450</td>
</tr>
<tr>
<td>#8</td>
<td>#1 and #7</td>
<td>4347</td>
</tr>
<tr>
<td>#9</td>
<td>&quot;randomized controlled trial&quot;:pt or &quot;randomized controlled trial as topic&quot;:pt or &quot;single-blind method&quot;:pt or &quot;double-blind method&quot;:pt or &quot;random allocation&quot;:pt</td>
<td>430710</td>
</tr>
<tr>
<td>#10</td>
<td>(review and systematic) or &quot;systematic review&quot; or (&quot;review literature as topic&quot; and systematic) or &quot;meta-analysis&quot;</td>
<td>64699</td>
</tr>
<tr>
<td>#11</td>
<td>[mh &quot;Cohort Studies&quot;] or [mh &quot;Epidemiologic Studies&quot;] or [mh &quot;Follow-up Studies&quot;] or &quot;prospective cohort&quot; or [mh &quot;prospective studies&quot;] or (prospective* and cohort and (study or studies))</td>
<td>154533</td>
</tr>
<tr>
<td>#12</td>
<td>#8 and #9 Publication Year from 2017</td>
<td>45</td>
</tr>
<tr>
<td>#13</td>
<td>#8 and #10 Publication Year from 2017</td>
<td>25</td>
</tr>
<tr>
<td>#14</td>
<td>#8 and #11 Publication Year from 2017</td>
<td>25</td>
</tr>
<tr>
<td>#15</td>
<td>#12 or #13 or #14</td>
<td>79 (78 imported)</td>
</tr>
</tbody>
</table>
Results: 82 imported after removing duplicates

The original search retrieved an original total of 184 records, and after initial removal of 40 duplicates, 144 records were left. Further removal of another 62 duplicates left 82 records for literature review.

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Last Run Via</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>&quot;rheumatoid arthritis&quot;</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4,779</td>
</tr>
<tr>
<td>S2</td>
<td>&quot;Adrenal Cortex Hormones&quot; OR corticosteroid</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>6,906</td>
</tr>
<tr>
<td>S3</td>
<td>Methylprednisolone OR prednisone OR prednisolone</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>6,715</td>
</tr>
<tr>
<td>S4</td>
<td>Hydroxychloroquine OR Lefunomide OR Methotrexate OR Sulfasalazine</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>5,371</td>
</tr>
<tr>
<td>S5</td>
<td>Adalimumab OR &quot;certolizumab pegol&quot; OR etanercept OR golimumab OR infliximab OR Abatacept OR tocilizumab OR rituximab OR Tofacitinib OR Sarilumab OR Baricitinib OR Sirukumab</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4,730</td>
</tr>
<tr>
<td>S6</td>
<td>amjevita OR Inflectra OR Erelzi</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2</td>
</tr>
<tr>
<td>S7</td>
<td>S2 OR S3 OR S4 OR S5 OR S6</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>20,004</td>
</tr>
<tr>
<td>S8</td>
<td>S1 AND S7</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2,493</td>
</tr>
<tr>
<td>#</td>
<td>Query</td>
<td>Limiters/Expanders</td>
<td>Last Run Via</td>
<td>Results</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>S9</td>
<td>S1 AND S7</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>899</td>
</tr>
<tr>
<td>S10</td>
<td>S1 AND S7</td>
<td>Limiters - Published Date: 20100101-20161231 Narrow by Language: English Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>878</td>
</tr>
<tr>
<td>S11</td>
<td>(&quot;Randomized Controlled Trial&quot; OR &quot;Single-Blind Method&quot; OR &quot;Double-Blind Method&quot; OR &quot;Random Allocation&quot; OR ((randomized OR randomised) AND (controlled AND trial)))</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4,755</td>
</tr>
<tr>
<td>S12</td>
<td>(&quot;review&quot; AND &quot;systematic&quot;) OR &quot;systematic review&quot; OR &quot;meta-analysis&quot;</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>3,143</td>
</tr>
<tr>
<td>S13</td>
<td>&quot;Case-Control Studies&quot; OR &quot;Cohort Studies&quot; OR &quot;Epidemiologic Studies&quot; OR &quot;Cross-Sectional Studies&quot; OR &quot;Organizational Case Studies&quot; OR &quot;Cross-Over Studies&quot; OR &quot;Follow-Up Studies&quot; OR &quot;Seroepidemiologic Studies&quot; OR &quot;Evaluation Studies&quot; OR &quot;observational study&quot; OR &quot;observational studies&quot;</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>1,628</td>
</tr>
<tr>
<td>S14</td>
<td>S10 AND S11</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>85</td>
</tr>
<tr>
<td>S15</td>
<td>S10 AND S12</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>64</td>
</tr>
<tr>
<td>S16</td>
<td>S10 AND S13</td>
<td>Limiters - Published Date: 20100101-20161231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>35</td>
</tr>
<tr>
<td>S17</td>
<td>S14 OR S15 OR S16</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>144</td>
</tr>
</tbody>
</table>
# Update Searches

**Results: 0 imported after removing duplicates**

The update search retrieved an original total of 11 records, and after the removal of all 11 as duplicates, no records were left for literature review.

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Last Run Via</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>&quot;rheumatoid arthritis&quot;</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4,867</td>
</tr>
<tr>
<td>S2</td>
<td>&quot;Adrenal Cortex Hormones&quot; OR corticosteroid</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>6,984</td>
</tr>
<tr>
<td>S3</td>
<td>Methylprednisolone OR prednisone OR prednisolone</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>6,788</td>
</tr>
<tr>
<td>S4</td>
<td>Hydroxychloroquine OR Leflunomide OR Methotrexate OR Sulfasalazine</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>5,459</td>
</tr>
<tr>
<td>S5</td>
<td>Adalimumab OR &quot;certolizumab pegol&quot; OR etanercept OR golimumab OR infliximab OR Abatacept OR tocilizumab OR rituximab OR Tofacitinib OR Sarilumab</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4,841</td>
</tr>
<tr>
<td>S6</td>
<td>amjevita OR Inflectra OR Erelzi</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>4</td>
</tr>
<tr>
<td>S7</td>
<td>S2 OR S3 OR S4 OR S5 OR S6</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>20,287</td>
</tr>
<tr>
<td>S8</td>
<td>S1 AND S7</td>
<td>Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2,536</td>
</tr>
<tr>
<td>S9</td>
<td>S1 AND S7</td>
<td>Narrow by Language: english Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2,427</td>
</tr>
<tr>
<td>#</td>
<td>Query</td>
<td>Limiters/Expanders</td>
<td>Last Run Via</td>
<td>Results</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------</td>
<td>---------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>S10</td>
<td>S1 AND S7</td>
<td>Limiters - Published Date: 20170101-20171231, Narrow by Language: English Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>36</td>
</tr>
<tr>
<td>S11</td>
<td>(&quot;Randomized Controlled Trial&quot; OR &quot;Single-Blind Method&quot; OR &quot;Double-Blind Method&quot; OR &quot;Random Allocation&quot; OR ((randomized OR randomised) AND controlled AND trial))</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>202</td>
</tr>
<tr>
<td>S12</td>
<td>(&quot;review&quot; AND &quot;systematic&quot;) OR &quot;systematic review&quot; OR &quot;meta-analysis&quot;</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>225</td>
</tr>
<tr>
<td>S13</td>
<td>&quot;Case-Control Studies&quot; OR &quot;Cohort Studies&quot; OR &quot;Epidemiologic Studies&quot; OR &quot;Cross-Sectional Studies&quot; OR &quot;Organizational Case Studies&quot; OR &quot;Cross-Over Studies&quot; OR &quot;Follow-Up Studies&quot; OR &quot;Seroepidemiologic Studies&quot; OR &quot;Evaluation Studies&quot; OR &quot;observational study&quot; OR &quot;observational studies&quot;</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>98</td>
</tr>
<tr>
<td>S14</td>
<td>S10 AND S11</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>7</td>
</tr>
<tr>
<td>S15</td>
<td>S10 AND S12</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2</td>
</tr>
<tr>
<td>S16</td>
<td>S10 AND S13</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>2</td>
</tr>
<tr>
<td>S17</td>
<td>S14 OR S15 OR S16</td>
<td>Limiters - Published Date: 20170101-20171231 Search modes - Boolean/Phrase</td>
<td>Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - International Pharmaceutical Abstracts</td>
<td>11</td>
</tr>
</tbody>
</table>
ClinicalTrials.gov:
April 12, 2017 Original and October 10, 2017 Updated Searches (Combined)

Results: 154
Completed Studies | Studies With Results | Rheumatoid Arthritis | "Adrenal Cortex Hormones"
OR corticosteroid OR Methylprednisolone OR prednisone OR prednisolone OR
Hydroxychloroquine OR Leflunomide OR Methotrexate OR Sulfasalazine OR Adalimumab OR
“certolizumab pegol” OR etanercept OR golimumab OR infliximab OR Abatacept OR
tocilizumab OR rituximab OR Tofacitinib OR Sarilumab OR amjevita OR Inflectra OR Erelzi |
Adult, Senior | First posted from 07/01/2010 to 10/05/2017

World Health Organization International Clinical Trials Registry Platform (WHO ICTRP):

April 12, 2017 Original Searches and October 10, 2017 Updated Searches (Combined)
Two searches, because of character limit:
For both:
Recruitment status: ALL
Date of registration between: July 1, 2010 and October 5, 2017

First search:
Results: 897 records for 394 trials; 394 (all) imported

Condition: Rheumatoid arthritis
Intervention: Adrenal Cortex Hormones OR corticosteroid OR Methylprednisolone OR prednisone OR prednisolone OR Hydroxychloroquine OR Leflunomide OR Methotrexate OR Sulfasalazine OR Adalimumab OR certolizumab pegol

Second search:
Results: 1205 records for 496 trials; 359 imported (137 duplicates)
Condition: Rheumatoid arthritis
Intervention: etanercept OR golimumab OR infliximab OR Abatacept OR tocilizumab OR rituximab OR Tofacitinib OR Sarilumab OR Baricitinib OR Sirukumab OR amjevita OR Inflectra OR Erelzi

New York Academy of Medicine Grey Literature Report:

April 12, 2017 Original Searches and October 5, 2017 Update Searches (Combined)
“rheumatoid arthritis”
Results: 5
Appendix B. Excluded Articles

X1 – Ineligible publication type
X2 – Population ages <19 yrs old
X3 – ≥50% patients have RA >2 yrs duration or non-RA diagnosis
X4 – Ineligible or no drug(s)
X5 – Ineligible or no comparator(s)
X6 – Ineligible or no outcome(s)
X7 – Ineligible treatment duration (<3 months of treatment)
X8 – Ineligible setting
X9 – Ineligible study design
X10 – Non-English language
X11 – Study protocol or abstract-only record (otherwise eligible)
X12 – Eligible except early RA up to 2 yrs
X13 – Excluded primary or companion article, to be cited in review
X14 – Irretrievable
X15 – Duplicate
X16 – Placebo-controlled study not usable in NWMA


B-41


<table>
<thead>
<tr>
<th>ID</th>
<th>Reference</th>
<th>Exclusion Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1172</td>
<td>Bijlsma JW. Disease control with glucocorticoid therapy in rheumatoid arthritis.</td>
<td>X9</td>
</tr>
</tbody>
</table>


1431. Combination Therapy with Adalimumab plus Methotrexate Significantly Improved Work Ability, Physical Function, Fatigue, and Other Patient-Reported Outcomes in Early Rheumatoid Arthritis: Results From a 26-Week Analysis. 2011 2011; MALDEN. WILEY-BLACKWELL; 63. Exclusion Code: X11.


1497. Graudal N, Juergens 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. p. 2852. Exclusion Code: X15.


## Appendix C. Detailed Evidence Table

### Appendix Table C-1. Evidence tables for randomized controlled trials and observational studies

<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Health Outcomes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
</table>
| **Author, yr, Study Name:** Atsumi et al., 2016-7 | Adults (aged 20-64) who are MTX naïve with RA fulfilling 2010 ACR/EULAR classification criteria, ≤ 12 months of persistent arthritic symptoms, DAS28-ESR ≥ 3.2, ≥ 3x upper limit of normal anti-CCP antibody, and positive rheumatoid factor and/or radiographic evidence of bone erosions | **Interventions, dose:**  
G1:  
- CZP: 400 mg at wks 0, 2, and 4, 200 mg every 2 wks thereafter (subcutaneous)  
- MTX: 8 mg/wk, increased to 12 mg/wk at wk 4, 16 mg/wk at wk 8, 16 mg/wk thereafter (oral)  
G2:  
- Placebo  
- MTX: 8 mg/wk, increased to 12 mg/wk at wk 4, 16 mg/wk at wk 8, 16 mg/wk thereafter (oral)  
Those in either arm with DAS28-ESR > 3.2 at/after wk 24 for ≥ 4 wks were eligible for rescue treatment with open-label CZP after discontinuing the double-blind period  
In the post-treatment period (wks 52-104) all patients received MTX alone. Patients who flared could receive rescue treatment with open-label CZP | **Mean disease duration, mos:**  
At 2 yrs  
DAS28-ESR LDA, %:  
G1: Figure only (Sup. Figure S1)  
G2: Figure only (Sup. Figure S1)  
P = 0.003  
Baseline DAS28-ESR, mean:  
5.4-5.5  
Baseline HAQ-DI, mean: Baseline HAQ-DI, mean:  
1.0-1.1  
Prior csDMARD use, %:  
18.5-19.5  
MTX naive:  
100  
MTX inadequate responders:  
NR  
Biologic non-responders:  
NR  
Prior CS use, %:  
16.4-19.7 | At 2 yrs  
Overall AEs:  
G1: 96.9  
G2: 95.5  
SAEs:  
G1: 10.7  
G2: 11.5  
Overall discontinuation:  
G1: 53.5  
G2: 63.7  
Discontinuation due to AEs:  
G1: 6.3  
G2: 3.8  
Discontinuation due to lack of efficacy:  
G1: 0.0  
G2: 0.6  
Patient adherence:  
NR  
Specific AEs:  
Deaths  
G1: 0.0  
G2: 0.0  
Malignancy  
G1: 1.3 (cervix carcinoma)  
G2: 0.0 | Medium (24 weeks); High (1-2 years) |
<p>| <strong>Country, Clinical Setting:</strong> Japan, multicenter | | | | | |
| <strong>Study Design:</strong> RCT | | | | | |
| <strong>Overall N:</strong> 316 | | | | | |
| <strong>Study Duration:</strong> 2 yrs | | | | | |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Atsumi et al., 2016-7&lt;sup&gt;13, 133&lt;/sup&gt;</td>
<td><strong>C-OPERA (continued)</strong></td>
<td>Mean age, yrs: 49 (range 21-64)</td>
<td><strong>No radiographic progression</strong> (change ≤ 0.5), %:</td>
<td></td>
<td><strong>Interstitial lung disease</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female: 81.0</td>
<td>G1: 84.2</td>
<td></td>
<td>G1: 4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td>G2: 67.5</td>
<td></td>
<td>G2: 0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: NR</td>
<td>P&lt;0.001</td>
<td></td>
<td><strong>Nausea/Vomiting/Decreased appetite</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF-36: NR</td>
<td><strong>At 1 year</strong></td>
<td></td>
<td>G1: 27.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS28-ESR disease activity: NR</td>
<td><strong>ACR20 response, %:</strong></td>
<td></td>
<td>G2: 24.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 78.6</td>
<td></td>
<td><strong>Hepatic disorders</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 68.8</td>
<td>p&lt;0.055</td>
<td>G1: 45.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>ACR50 response, %:</strong></td>
<td></td>
<td>G2: 46.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 73.0</td>
<td>p&lt;0.001</td>
<td><strong>Tuberculosis</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 51.6</td>
<td></td>
<td>G1: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td></td>
<td>G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>ACR70 response, %:</strong></td>
<td></td>
<td><strong>Pneumonia</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 57.2</td>
<td>p&lt;0.001</td>
<td>G1: 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 34.4</td>
<td></td>
<td>G2: 6.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td></td>
<td><strong>Serious Infections</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>DAS28-ESR remission, %:</strong></td>
<td></td>
<td>G1: 3.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 57.2</td>
<td></td>
<td>G2: 5.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 36.9</td>
<td></td>
<td><strong>Infections and infestations</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td></td>
<td>G1: 71.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>mTSS score</strong></td>
<td></td>
<td>G2: 59.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change from baseline, mean: G1: 0.36 (SD, 2.70)</td>
<td></td>
<td></td>
<td><strong>Injection site reaction</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 1.58 (SD, 4.86)</td>
<td></td>
<td></td>
<td>G1: 3.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td></td>
<td></td>
<td>G2: 1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>At 1 yr</strong></td>
<td></td>
<td></td>
<td><strong>Overall AEs:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall AEs:</td>
<td></td>
<td></td>
<td>G1: 96.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 8.2</td>
<td></td>
<td></td>
<td>G2: 94.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 8.9</td>
<td></td>
<td></td>
<td><strong>SAEs:</strong></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomesa</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Atsumi et al., 2016-7</td>
<td><strong>C-OPERA</strong> (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No radiographic progression (change ≤ 0.5), %:</td>
<td></td>
<td>Overall discontinuation:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 82.9</td>
<td>G1: 30.2</td>
<td>G2: 53.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 70.7</td>
<td>Discontinuation due to AEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p=0.011</td>
<td>G1: 5.7</td>
<td>G2: 3.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to lack of efficacy:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HAQ:</td>
<td>G1: 0.0</td>
<td>G2: 0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td>Patient adherence:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific AEs:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Deaths</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.0</td>
<td>G2: 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Malignancy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.6 (cervix carcinoma)</td>
<td>G2: 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Interstitial lung disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.1</td>
<td>G2: 0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nausea/Vomiting/Decreased appetite</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 24.5</td>
<td>G2: 20.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hepatic disorders</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 42.8</td>
<td>G2: 44.6</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Atsumi et al., 2016-7</td>
<td>C-OPERA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tuberculosis</td>
<td>G1: 0.0&lt;br&gt;G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pneumonia</td>
<td>G1: 4.4&lt;br&gt;G2: 5.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Serious Infections</td>
<td>G1: 3.1&lt;br&gt;G2: 4.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Infections and infestations</td>
<td>G1: 61.0&lt;br&gt;G2: 55.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Injection site reaction</td>
<td>G1: 3.1&lt;br&gt;G2: 1.3</td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics

**Author, yr, Study Name:** Atsumi et al., 2016

**Study Population Summary**

- Ethnicity, % Latino: NR

**Interventions and Patient Characteristics**

- RF seropositive, %: 93.0-96.2
- anti-CCP seropositive, %: 100
- Baseline mTSS score, mean: 5.2-6.0
- Bone erosion judged by physician, %: 49.7-51.0

**Baseline Disease and Treatment Characteristics**

- SF-36: NR
- At 24 wks DAS28-ESR remission, %:
  - G1: 52.8
  - G2: 30.6
  - p<0.001
- mTSS score, mean change from baseline:
  - G1: 0.26 (SD, 1.55)
  - G2: 0.86 (SD, 2.37)
  - p=0.003

**Health Outcomes**

**Adverse Events (%)**

- Hepatic disorders
  - G1: 42.8
  - G2: 44.6
- Tuberculosis
  - G1: 0.0
  - G2: 0.0
- Pneumonia
  - G1: 4.4
  - G2: 5.1
- Serious Infections
  - G1: 3.1
  - G2: 4.5
- Infections and infestations
  - G1: 61.0
  - G2: 55.4
- Injection site reaction
  - G1: 3.1
  - G2: 1.3

**ROB Rating**
### Study Characteristics

<table>
<thead>
<tr>
<th>Author, yr, Study Name: Bakker et al., 2012; CAMERA-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country, Clinical Setting: Netherlands, 7 hospital outpatient rheumatology clinics</td>
</tr>
<tr>
<td>Study Design: RCT</td>
</tr>
<tr>
<td>Overall N: 239</td>
</tr>
<tr>
<td>Study Duration: 2 years</td>
</tr>
</tbody>
</table>

### Study Population Summary

Patients meeting 1987 revised ACR criteria for RA with disease duration <1 yr, who were DMARD and glucocorticoid naive.

### Interventions and Patient Characteristics

<table>
<thead>
<tr>
<th>Interventions, dose:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: MTX + Prednisone 10 mg/d tight-control strategy</td>
</tr>
<tr>
<td>G2: MTX + Placebo tight-control strategy</td>
</tr>
<tr>
<td>Both arms received initial dose of oral MTX 10 mg/wk, plus folic acid 0.5 mg/d, bisphosphonate (alendronate or risedronate) and cholecalciferol.</td>
</tr>
<tr>
<td>Strategy steps based on &gt;20% improvement in SJC and at least 2 of the following: TJC, ESR, and VAS for general well-being at each monthly visit, compared with previous visit.</td>
</tr>
<tr>
<td>Steps to achieve &gt;20% improvement could include MTX dose escalation, switch to subcutaneous MTX, addition of cyclosporine or adalimumab, or switch to different medication (the latter leading to dropout)</td>
</tr>
</tbody>
</table>

| G1: 118 |
| G2: 121 |

### Mean age, yrs: |

53-54 |

<table>
<thead>
<tr>
<th>Sex, % female:</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-61</td>
</tr>
</tbody>
</table>

### Baseline Disease and Treatment Characteristics

<table>
<thead>
<tr>
<th>Mean disease duration, mos: NR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline DAS, mean: 5.5-5.8</td>
</tr>
<tr>
<td>Baseline HAQ, mean: 1.0-1.2</td>
</tr>
<tr>
<td>MTX naive: 100</td>
</tr>
<tr>
<td>Prior csDMARD use, %: 0</td>
</tr>
<tr>
<td>MTX inadequate responders: NA</td>
</tr>
<tr>
<td>Biologic non-responders: NA</td>
</tr>
<tr>
<td>Seropositive (RF or CCP) (%): RF+: 55-61</td>
</tr>
<tr>
<td>Baseline Sharp score, median: 0</td>
</tr>
<tr>
<td>Erosive disease, %: 12-17</td>
</tr>
</tbody>
</table>

### Health Outcomes

<table>
<thead>
<tr>
<th>At 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean DAS28 score (SD)</td>
</tr>
<tr>
<td>G1: 2.30 (0.34)</td>
</tr>
<tr>
<td>G2: 2.49 (0.25)</td>
</tr>
<tr>
<td>Mean difference (95% CI): -0.26 (-0.68 to 0.16) (p=0.21)</td>
</tr>
<tr>
<td>ACR20 response, %</td>
</tr>
<tr>
<td>G1: 65</td>
</tr>
<tr>
<td>G2: 61</td>
</tr>
<tr>
<td>Mean difference (95% CI): 3.6 (-8.7 to 15.9) (p=0.56)</td>
</tr>
<tr>
<td>ACR50 response, %</td>
</tr>
<tr>
<td>G1: 53</td>
</tr>
<tr>
<td>G2: 42</td>
</tr>
<tr>
<td>Mean difference (95% CI): 11.0 (-1.7 to 23.6) (p=0.091)</td>
</tr>
<tr>
<td>ACR70 response, %</td>
</tr>
<tr>
<td>G1: 38</td>
</tr>
<tr>
<td>G2: 19</td>
</tr>
<tr>
<td>Mean difference (95% CI): 18.3 (7.0 to 29.6) (p=0.002)</td>
</tr>
<tr>
<td>Remission, %</td>
</tr>
<tr>
<td>G1: 72</td>
</tr>
<tr>
<td>G2: 61</td>
</tr>
<tr>
<td>Mean difference (95% CI): 10.5 (-1.5 to 22.4) (p=0.089)</td>
</tr>
<tr>
<td>Median total SHS score (IQR)</td>
</tr>
<tr>
<td>G1: 0 (0 to 3)</td>
</tr>
<tr>
<td>G2: 0 (0 to 4) (p=0.32)</td>
</tr>
<tr>
<td>Mean difference (95% CI): 0.0 (-1.1 to 1.1)</td>
</tr>
</tbody>
</table>

### Adverse Events (%)

| Overall: |
| G1: 74 |
| G2: 79 |
| SAEs |
| G1: 2 |
| G2: 4 |

### Other Events

<table>
<thead>
<tr>
<th>Overall discontinuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>At 2 years</td>
</tr>
<tr>
<td>G1: 28.0</td>
</tr>
<tr>
<td>G2: 29.8</td>
</tr>
<tr>
<td>Discontinuation because of AEs</td>
</tr>
<tr>
<td>At 2 years</td>
</tr>
<tr>
<td>G1: 13.6</td>
</tr>
<tr>
<td>G2: 16.5</td>
</tr>
<tr>
<td>At 1 year</td>
</tr>
<tr>
<td>G1: 8.5</td>
</tr>
<tr>
<td>G2: 7.4</td>
</tr>
</tbody>
</table>

### Patient adherence

| At 2 years |
| G1: 94.9 |
| G2: 96.6 |
| At 1 year |
| G1: 95.7 |
| G2: 97.5 |

### ROB Rating

Medium
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Bakker et al., 2012; CAMERA-II (continued)</td>
<td>Race, % white: NR</td>
<td>Sensitivity analyses for observed data showed no statistical differences</td>
<td><strong>Median SHS erosive joint damage score (IQR)</strong></td>
<td>G1: 0 (0 to 0) G2: 0 (0 to 2) (P =0.022)</td>
<td><strong>Mortality</strong></td>
<td>G1: 1 G2: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean difference (95% CI): 0.0 (-0.1 to 0.0)</td>
<td>Linear mixed-model analysis found that erosion score was, on average, 0.87 SHS units lower in G1 than G2</td>
<td><strong>Hospitalization</strong></td>
<td>G1: 1 G2: 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Linear mixed-model regression coefficient (95% CI): -0.87 (-1.31 to -0.43) (p=0.001)</td>
<td><strong>Nausea</strong></td>
<td>G1: 19.65 G2: 36.1 (p=0.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosion-free as measured by SHS, %</td>
<td>G1: 78 G2: 67 (p=NR)</td>
<td><strong>ALT &gt;ULN</strong></td>
<td>G1: 12.8 G2: 27.7 (p=0.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Mean HAQ score (SD)</strong></td>
<td>G1: 0.5 (0.13) G2: 0.7 (0.13)</td>
<td><strong>AST &gt;ULN</strong></td>
<td>G1: 6.8 G2: 17.6 (p=0.016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean difference (95% CI): -0.18 (-0.34 to -0.02) (p=0.027)</td>
<td><strong>Infections requiring antibiotics</strong></td>
<td>G1: 0.01 G2: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>At 18 months</td>
<td><strong>Pneumonitis</strong></td>
<td>G1: 0.01 G2: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean DAS28 score (SD)</td>
<td>Figure only data (p=0.183)</td>
<td><strong>Headache</strong></td>
<td>G1: 19.6 G2: 26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean HAQ score (SD)</td>
<td>Figure only data; p=0.014</td>
<td><strong>Weight gain (kg, mean [SD])</strong></td>
<td>G1: 2.9 (4.2) G2: 1.3 (5.3) (p=0.028)</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Bakker et al., 2012; CAMERA-II (continued)</td>
<td></td>
<td></td>
<td></td>
<td>At 1 year</td>
<td>Mean DAS28 score (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2.45 (0.29)</td>
<td>G2: 2.59 (0.29)</td>
<td>Mean difference (95% CI): -0.21 (-0.52 to 0.11) (p=0.194)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td>G1: 70</td>
<td>G2: 66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td>G1: 56</td>
<td>G2: 43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td>G1: 27</td>
<td>G2: 26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean HAQ score (SD)</td>
<td>G1: 0.5 (0.11)</td>
<td>G2: 0.7 (0.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 6 months</td>
<td>Mean DAS28 score (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean difference (95% CI): -0.89 (-0.52 to -0.11) (p&lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean HAQ score (SD)</td>
<td>Figure only data (p=0.001)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Bakker et al., 2012; ÇCAMERA-II (continued)</td>
<td></td>
<td></td>
<td></td>
<td>At 3 months</td>
<td>Mean DAS28 score (SD)</td>
<td>-1.56 (-1.88 to -1.25) (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean difference (95% CI):</td>
<td>Mean HAQ score (SD)</td>
<td>Figure only data (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td>At year 2 (open-label extension)</td>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>Batson et al.,</td>
<td></td>
<td></td>
<td></td>
<td>ACR20, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200014</td>
<td></td>
<td></td>
<td></td>
<td>G1: 59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G2: 72 (p=0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002100</td>
<td></td>
<td></td>
<td></td>
<td>ACR50, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G1: 42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005112</td>
<td></td>
<td></td>
<td></td>
<td>G2: 49 (p=NS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batson et al.,</td>
<td></td>
<td></td>
<td></td>
<td>ACR70, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006111</td>
<td></td>
<td></td>
<td></td>
<td>G1: 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enbrel ERA</td>
<td></td>
<td></td>
<td></td>
<td>G2: 29 (p=NS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Setting:</td>
<td></td>
<td></td>
<td></td>
<td>Subgroup analysis for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US, clinics</td>
<td></td>
<td></td>
<td></td>
<td>ACR20/50/70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design:</td>
<td></td>
<td></td>
<td></td>
<td>Ages ≥65, events per patient-year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td>ACR20, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td>424 eligible (of 632 total)</td>
<td></td>
<td></td>
<td>G1: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>12 mos (1 year open label extension)</td>
<td></td>
<td></td>
<td>G2: 54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs:</td>
<td>49-51</td>
<td></td>
<td></td>
<td>Prior CS use, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td>74-75</td>
<td></td>
<td></td>
<td>23-25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td>84-88</td>
<td></td>
<td></td>
<td>MTX naïve, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Sharp score,</td>
<td>2.4-12.9</td>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean disease duration,</td>
<td>11-12</td>
<td></td>
<td></td>
<td>Prior csDMARD use, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mos:</td>
<td>11-12</td>
<td></td>
<td></td>
<td>23-25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median disease duration,</td>
<td>0.3-0.8</td>
<td></td>
<td></td>
<td>39-42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mos:</td>
<td>0.3-0.8</td>
<td></td>
<td></td>
<td>MTX inadequate responders, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interventions, dose:</td>
<td>G1: MTX 7.5 mg/wk to 20 mg/wk dose escalation (19 mg/wk mean dose)</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: ETN (25 mg twice wkly, subcutaneous)</td>
<td></td>
<td></td>
<td></td>
<td>Biologic non-responders, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>G1: 217</td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 207</td>
<td></td>
<td></td>
<td></td>
<td>Seropositive (RF or CCP) (%):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF+:</td>
<td>87-89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At year 2</td>
<td>Overall discontinuation</td>
<td>G1: 40.55</td>
<td>G2: 25.6 (p=NR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinuation because of AEs</td>
<td>G1: 12.4</td>
<td>G2: 7.25 (p=NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinuation because of lack of efficacy</td>
<td>G1: 3.7</td>
<td>G2: 4.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient adherence</td>
<td>SAEs</td>
<td>G1: 12</td>
<td>G2: 12 (p=NR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subgroup analysis for SAEs</td>
<td>Ages ≥65, events per patient-year</td>
<td>G1: 0.417</td>
<td>G2: 0.321</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ages &lt;65, events per patient-year</td>
<td>G1: 0.072</td>
<td>G2: 0.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rates similar in elderly vs. non-elderly patients, but P=NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 200014;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 200210;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 200512;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 200611</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enbrel ERA</td>
<td></td>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Baseline DAS, mean:** NR
- **Erosive disease, %:** 85-88

**Baseline Disease and Treatment Characteristics**

- **ACR70, %**
  - G1: 25
  - G2: 32

- **HAQ improvement of at least 0.5 units, %**
  - G1: 37
  - G2: 55
  - G2 > G1 (p<0.001)

**Subgroup analysis for mean change in HAQ from baseline (SD)**

- Ages ≥65, events per patient-year
  - G1: 0.61 (0.78)
  - G2: 0.46 (0.66)
  - Both groups showed improvements exceeding MCID

- Ages <65, events per patient-year
  - NR, but improvements mirrored those of ages ≥65

- **Change in total modified Sharp score, mean**
  - G1: 3.2
  - G2: 1.3 (p=0.001)

- **Erosion score change, mean**
  - G1: 1.9
  - G2: 0.7 (p=0.001)

**Health Outcomes**

- **Mortality**
  - G1: 0
  - G2: 1 (p=NR)

- **Serious infections**
  - G1: 4.15
  - G2: 3.4 (p=NR)

- **Subgroup analysis for serious infections**
  - Ages ≥65, events per patient-year
    - G1: 0.074
    - G2: 0.095
  - Ages <65, events per patient-year
    - G1: 0.016
    - G2: 0.01

- **Rates higher in elderly patients, but P=NR**

- **Injection site reaction**
  - G1: 9
  - G2: 39
  - G1 < G2 (p ≤0.05)

- **Nausea**
  - G1: 31
  - G2: 20
  - G1 > G2 (p ≤0.05)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batson et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batson et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200611</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enbrel ERA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**At 1 yr**

ACR20 response rates, %:
- G1: 65
- G2: 72 (p=0.16)

Increase in Sharp score, mean:
- G1: 1.59
- G2: 1.00 (p=0.11)

Erosion score change, mean:
- G1: 0.47
- G2: 1.03 (p=0.002)

Mean HAQ scores:
No significant difference in HAQ scores between MTX and ETN 25 mg arms, with ~55% in each arm having at least a 0.5-unit improvement.

**At 6 months**

Significantly more pts on ETN (25 mg) than on MTX achieved ACR20, ACR50, ACR70 responses (data NR, p<0.05)

**Adverse Events (%)**

<table>
<thead>
<tr>
<th>Event</th>
<th>G1</th>
<th>G2</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dizziness</td>
<td>12</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Vomiting</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Alopecia</td>
<td>12</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mouth ulcer</td>
<td>17</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>3</td>
<td>4</td>
<td>p=NR</td>
</tr>
</tbody>
</table>

**Subgroup analysis for cancer**

Ages ≥65, events per patient-year:
- G1: 0.049
- G2: 0.057
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enbrel ERA (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Ages <65, events per patient-year
  - G1: 0.004
  - G2: 0.003

- At year 1
  - Overall discontinuation
    - G1: 22.1
    - G2: 14.5 (p=NR)

- Discontinuation because of AEs
  - G1: 10.1
  - G2: 4.8 (p=NR)

- Patient adherence
  - NR

- Mortality
  - G1: 0
  - G2: 1

- URTI
  - G1: 39
  - G2: 35
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al., 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathon et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enbrel ERA (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Infections at other respiratory tract sites, events per patient-year
  - G1: 1.3
  - G2: 1.0 (p=0.006)

- Injection site reaction
  - G1: 7
  - G2: 37
  - G1 < G2 (p <0.05)

- Nausea
  - G1: 29
  - G2: 17
  - G1 > G2 (p<0.05)

- Rash
  - G1: 23
  - G2: 12
  - G1 > G2 (p <0.05)

- Alopecia
  - G1: 12
  - G2: 6
  - G1 > G2 (p <0.05)

- Mouth ulcer
  - G1: 14
  - G2: 5
  - G1 > G2 (p<0.05)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>Bejarano et al., 2008; Emery et al., 2016</td>
<td>Patients aged ≥18 years with RA according to ACR criteria, &lt;2 yrs symptom duration, MTX/biologic naïve, who were in paid employment, and had self-reported RA-related work impairment</td>
<td>Interventions, dose: G1: ADA (40 mg every other wk + MTX (7.5 mg/wk, max 25 mg/wk) G2: Placebo + MTX (7.5 mg/wk, max 25 mg/wk) MTX: Dosage reached 25 mg/wk by wk 12 in the presence of remaining synovitis. Mean dose at 56 wks was 15.5 mg/wk in G1 and 16.2 mg/wk in G2 ADA: Administered via subcutaneous injection</td>
<td>Mean symptom duration, mos: 7.9-9.5</td>
<td>Mean symptom duration, mos: 7.9-9.5</td>
<td>Mean symptom duration, mos: 7.9-9.5</td>
</tr>
<tr>
<td>Overall N:</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>56 wks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td>United Kingdom, “Multicenter”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design:</td>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Disease and Treatment Characteristics</td>
<td></td>
<td></td>
<td>Health Outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age (SD), yrs:</td>
<td>47 (SD, 9.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td>53.4-58.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interventions, dose:</td>
<td>G1: ADA (40 mg every other wk + MTX (7.5 mg/wk, max 25 mg/wk) G2: Placebo + MTX (7.5 mg/wk, max 25 mg/wk) MTX: Dosage reached 25 mg/wk by wk 12 in the presence of remaining synovitis. Mean dose at 56 wks was 15.5 mg/wk in G1 and 16.2 mg/wk in G2 ADA: Administered via subcutaneous injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean symptom duration, mos:</td>
<td>7.9-9.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline DAS, mean:</td>
<td>5.9-6.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline HAQ, mean (SD):</td>
<td>1.3 (SD, 0.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX naïve, %:</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior csDMARD use:</td>
<td>Mean: 0.2 per patient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX inadequate responders:</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologic non-responders:</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF seropositive (%):</td>
<td>95-96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-CCP antibody positive (%):</td>
<td>63-64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Sharp score, mean:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erosive disease, %:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At week 56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity</td>
<td>G1: 3.0 (SD, 1.8) G2: 3.8 (SD, 2.1, p=0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20 response, %</td>
<td>G1: 71.6 G2: 54.8 (p=0.034)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR50 response, %</td>
<td>G1: 56.0 G2: 45.2 (p=0.189)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR70 response, %</td>
<td>G1: 50.7 G2: 37.5 (p=0.108)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remission (DAS28 &lt;2.6), %</td>
<td>G1: 48.0 G2: 36.1 (p=0.145)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHS</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ change from baseline</td>
<td>G1: -0.7 (SD, 0.6) G2: -0.4 (SD, 0.7) (p=0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job loss, %</td>
<td>G1: 18.6 G2: 39.7 (p&lt;0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job loss, %</td>
<td>G1: 16 G2: 27.3 (p=0.092)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinuation because of AEs</td>
<td>G1: 8 G2: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinuation because of lack of efficacy</td>
<td>G1: 17.3 G2: 35.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient adherence</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain (Serious)</td>
<td>G1: 1.4 G2: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td>G1: 21.3 G2: 32.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diarrhea</td>
<td>G1: 10.7 G2: 8.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>G1: 10.7 G2: 6.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Bijlsma et al., 2016;33</td>
<td>Patients were diagnosed with RA within 1 year before inclusion, DMARD-naive, aged ≥18, met current RA classification criteria, and had a DAS28 score of ≥2.6</td>
<td>Interventions, dose: G1: TCZ 8 mg/kg intravenously every 4 wks (max 800 mg/dose) + MTX 10 mg/wk orally (max 30 mg/wk) G2: TCZ 8 mg/km intravenously every 4 wks (max 800 mg/dose) + placebo MTX G3: MTX 10 mg/wk orally (max 30 mg/wk) + placebo TCZ</td>
<td>Median disease duration, days (IQR): 26 (IQR, 16.0-43.0) Baseline DAS, mean: 5.2 (SD, 1.1) Baseline HAQ, mean: 1.2 (SD, 0.64) MTX naïve, %: 100 Prior csDMARD use, %: 0 MTX inadequate responders: NA Biologic non-responders, %: NA</td>
<td>At 2 yrs DAS disease activity, decrease from baseline, median (min, max) G1: 3.3 (-0.73, 6.07) G2: 3.3 (0.1, 6.8) G3: 3.2 (-0.79, 7.52) p=0.66 ACR20 response, % G1: 63 G2: 65 G3: 61 ACR50 response, % G1: 49 G2: 55 G3: 48 ACR70 response, % G1: 36 G2: 39 G3: 35</td>
<td>Overall: G1: 99.1 G2: 96.1 G3: 98.1 p=0.32 SAEs G1: 16.0 G2: 18.4 G3: 12.0 p=0.44 Overall discontinuation G1: 26.4 G2: 21.4 G3: 27.8 Discontinuation because of AEs G1: 8.5 G2: 9.7 G3: 7.4 p=0.82 Discontinuation because of lack of efficacy G1: 8.5 G2: 3.9 G3: 12.0</td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td>The Netherlands, Outpatient departments</td>
<td>Study Design: RCT</td>
<td>Overall N: 317</td>
<td>Study Duration: 2 yrs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C-16
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td>Sex, % female: 67</td>
<td>Seropositive (RF or anti-CCP) (%): RF: 72, anti-CCP: 70, Combined RF and anti-CCP: 79</td>
<td>DAS remission, %, sustained during entire study</td>
<td>Patient adherence, %</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: 96</td>
<td></td>
<td>G1: 86</td>
<td>G1: 73.6</td>
<td>33</td>
</tr>
<tr>
<td>U-Act-Early (continued)</td>
<td></td>
<td></td>
<td>Baseline Sharp score, median (IQR): 0.0 (IQR, 0.0-1.0)</td>
<td>G2: 88</td>
<td>G2: 78.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosive disease, %: NR</td>
<td>G3: 77</td>
<td>G3: 72.2</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SHS change from baseline, mean (SD)</td>
<td>G2 vs. G3: p=0.0356</td>
<td>G2: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 1.18 (SD, 3.919)</td>
<td>G2 vs. G1: p=0.59</td>
<td>G3: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 1.45 (SD, 4.272)</td>
<td>G3 vs. G1: p=0.06</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 1.53 (SD, 2.421)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>median (IQR) G1: 0.00 (IQR, 0.00-1.00)</td>
<td>G2 vs. G3: p=0.0381</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 0.00 (IQR, 0.00-2.00)</td>
<td>G2 vs. G1: p=0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 0.00 (IQR, 0.00-2.56)</td>
<td>G3 v G1: p=0.0207</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HAQ mean change from baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 0.48 (SD, 0.55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 0.61 (SD, 0.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 0.62 (SD, 0.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p=0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FACIT-F Score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 39.2 (10.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 38.4 (10.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 37.9 (10.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FACIT-F change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 6.3 (10.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 6.4 (11.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 6.6 (11.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 64.9 (15.5)</td>
<td>G1: 64.9 (15.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 65.0 (19.1)</td>
<td>G2: 65.0 (19.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 63.6 (16.4)</td>
<td>G3: 63.6 (16.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 15.3 (18.9)</td>
<td>G1: 15.3 (18.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 16.0 (19.1)</td>
<td>G2: 16.0 (19.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 73.5 (14.3)</td>
<td>G1: 73.5 (14.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 73.6 (15.2)</td>
<td>G2: 73.6 (15.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 73.2 (15.2)</td>
<td>G3: 73.2 (15.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 11.6 (12.6)</td>
<td>G1: 11.6 (12.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 6.9 (16.1)</td>
<td>G2: 6.9 (16.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 8.6 (16.2)</td>
<td>G3: 8.6 (16.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D Score, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.85 (0.15)</td>
<td>G1: 0.85 (0.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.83 (0.19)</td>
<td>G2: 0.83 (0.19)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.82 (0.15)</td>
<td>G3: 0.82 (0.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.18 (0.24)</td>
<td>G1: 0.18 (0.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.22 (0.27)</td>
<td>G2: 0.22 (0.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.16 (0.22)</td>
<td>G3: 0.16 (0.22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-VAS, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 73.8 (16.7)</td>
<td>G1: 73.8 (16.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 76.6 (18.2)</td>
<td>G2: 76.6 (18.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 71.2 (17.3)</td>
<td>G3: 71.2 (17.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-VAS change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 12.6 (20.6)</td>
<td>G1: 12.6 (20.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 14.0 (21.0)</td>
<td>G2: 14.0 (21.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 12.3 (21.4)</td>
<td>G3: 12.3 (21.4)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td>U-Act-Early (continued)</td>
<td></td>
<td></td>
<td>At 52 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS disease activity, decrease from baseline, median (min, max)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.3 (-1.02, 7.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 3.4 (0.28, 7.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 3.3 (-0.74, 6.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 72</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 69</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 51</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS change from baseline, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.50 (SD, 1.495)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.79 (SD, 3.242)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.96 (SD, 2.870)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.00 (IQR, 0.00-0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.00 (IQR, 0.00-0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.00 (IQR, 0.00-0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G3: p=0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G1: p=0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3 v G1: p=0.0164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>1. <strong>Author, yr, Study Name:</strong> Bijlsma et al., 2016; Teitsma et al., 2017.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. <strong>U-Act-Early</strong> (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. <strong>Study Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. <strong>HAQ mean change from baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.46 (SD, 0.50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0.48 (SD, 0.55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 0.55 (SD, 0.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. <strong>FACIT-F Score, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 39.5 (8.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 40.5 (9.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 37.1 (11.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FACIT-F change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 7.2 (9.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 8.0 (11.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 6.7 (12.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 PCS, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 68.5 (16.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 70.5 (19.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 64.2 (17.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2 vs. G3, P&lt;0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 PCS change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 19.2 (16.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 21.5 (17.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 14.3 (16.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2 vs. G3, P&lt;0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 MCS, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 74.7 (13.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 76.0 (14.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 73.8 (15.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 MCS change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 12.0 (13.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 11.8 (16.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 10.7 (16.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D Score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.89 (0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0.84 (0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 0.84 (0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Name: Bijlsma et al., 2016;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teitsma et al., 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-Act-Early</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.20 (0.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0.22 (0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 0.20 (0.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-VAS, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 74.0 (16.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 74.1 (18.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 72.6 (20.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-VAS change from BL, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 13.8 (23.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 12.6 (19.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 14.2 (21.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 24 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity, median decrease from baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 3.6 (0.75, 7.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 3.6 (0.45, 7.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 2.1 (-1.67, 5.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p &lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2 vs. G3: p=0.0343</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3 vs. G1: p=0.0099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR50 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2 vs. G3: p=0.0009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3 vs. G1: p&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-Act-Early (continued)</td>
<td></td>
<td></td>
<td></td>
<td>G1: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G3: p=0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3 vs. G1: P &lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS change from baseline, mean (SD)/median IQR</td>
<td>G1: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean change from baseline</td>
<td>G1: 0.50 (SD, 0.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.63 (SD, 0.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.65 (SD, 0.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.0275</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FACIT-F Score, mean (SD)</td>
<td>G1: 39.1 (9.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 39.0 (9.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 36.0 (8.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 and G2 vs. G3, p=0.038 each</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FACIT-F change from BL, mean (SD)</td>
<td>G1: 7.4 (9.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 7.3 (10.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 4.7 (9.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 and G2 vs. G3, p&lt;0.05 each</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS, mean (SD)</td>
<td>G1: 64.9 (18.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 63.0 (18.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 60.2 (16.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr,</strong></td>
<td><strong>Study Name:</strong> Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td><strong>U-Act-Early</strong> (continued)</td>
<td><strong>Interventions and Patient Characteristics</strong></td>
<td><strong>Baseline Disease and Treatment Characteristics</strong></td>
<td><strong>Health Outcomes</strong></td>
<td><strong>Adverse Events (%)</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS change from BL, mean (SD)</td>
<td>G1: 15.6 (16.8)</td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS, mean (SD)</td>
<td>G1: 73.8 (16.4)</td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS change from BL, mean (SD)</td>
<td>G1: 10.3 (15.0)</td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-5D Score, mean (SD)</td>
<td>G1: 0.84 (0.17)</td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-5D change from BL, mean (SD)</td>
<td>G1: 0.19 (0.22)</td>
<td>G1 vs. G3, p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-VAS, mean (SD)</td>
<td>G1: 72.8 (19.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Bijlsma et al., 2016; Teitsma et al., 2017</td>
<td>U-Act-Early (continued)</td>
<td></td>
<td></td>
<td>EQ-VAS change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 12.6 (21.9)</td>
<td>G2: 11.7 (20.5)</td>
<td>G3: 10.7 (20.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FACIT-F Score, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 35.7 (10.8)</td>
<td>G2: 38.2 (9.4)</td>
<td>G3: 35.1 (10.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FACIT-F change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 4.8 (9.0)</td>
<td>G2: 6.0 (10.5)</td>
<td>G3: 4.0 (10.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 59.7 (18.7)</td>
<td>G2: 61.6 (15.8)</td>
<td>G3: 57.6 (15.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS change from BL, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 10.2 (13.9)</td>
<td>G2: 13.6 (14.8)</td>
<td>G3: 6.6 (12.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 70.1 (15.8)</td>
<td>G2: 72.7 (14.7)</td>
<td>G3: 69.4 (14.1)</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Bijlsma et al., 2016; 33 Teitsma et al., 2017 135</td>
<td>U-Act-Early (continued)</td>
<td>SF-36 MCS change from BL, mean (SD)</td>
<td>G1: 7.6 (13.6)</td>
<td>G2: 7.3 (13.7)</td>
<td>G3: 4.7 (13.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-5D Score, mean (SD)</td>
<td>G1: 0.79 (0.20)</td>
<td>G2: 0.80 (0.14)</td>
<td>G3: 0.74 (0.21)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3, p=0.041</td>
<td>G2 vs. G3, p=0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-5D change from BL, mean (SD)</td>
<td>G1: 0.14 (0.24)</td>
<td>G2: 0.18 (0.24)</td>
<td>G3: 0.08 (0.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1 and G2 vs. G3, p&lt;0.05 each</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-VAS, mean (SD)</td>
<td>G1: 72.4 (15.9)</td>
<td>G2: 69.5 (16.4)</td>
<td>G3: 63.9 (17.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3, p=0.001</td>
<td>G2 vs. G3, p=0.039</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EQ-VAS change from BL, mean (SD)</td>
<td>G1: 12.6 (20.8)</td>
<td>G2: 8.8 (19.0)</td>
<td>G3: 3.9 (19.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1 and G2 vs. G3, p&lt;0.05 each</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Bili et al., 2014¹¹</td>
<td>Patients had RA diagnosis made with the International Classification of Diseases, Ninth Revision (ICD-9), twice by a GHS rheumatologist (definition was validated against 1987 ACR criteria). Patients with CVD prior to initiation of RA medication, and those who were DMARD naïve were excluded</td>
<td><strong>Interventions, dose:</strong> G1: TNFa inhibitors alone or in combination with MTX medication exposure G2: MTX alone or in combination with other nonbiologic DMARDs G3: Non-MTX, nonbiologic DMARDs G1 details: TNFa inhibitors include: ETN, ADA, IFX, GOL, and certolizumab. Other concomitant nonbiologic DMARDs permitted G2 details: Nonbiologic DMARDs include: MTX, HCQ, LEF, Azathioprine, SSZ, and Minocycline. Could not also use TNFa inhibitors or other biologic medicines Note: in all groups, Corticosteroids were considered non-DMARDs and (along with NSAIDs) were allowed in each group. Dose information for all groups not available. Additionally, Patients could contribute time to different groups according to medication exposure. Therefore, exposure is reported as “exposure periods” and one patient can contribute to multiple periods</td>
<td><strong>Median disease duration, mos (IQR):</strong> 0.99-9.0 mos</td>
<td><strong>Baseline DAS, mean:</strong> NR</td>
<td><strong>Baseline HAQ, median:</strong> NR</td>
<td><strong>MTX naive:</strong> NR</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Bili et al., 2014¹¹ (continued)</td>
<td>N:</td>
<td>Baseline Sharp score, mean: NR</td>
<td>CVD hazard ratio, fully adjusted, (95% CI)</td>
<td>G1: 0.79 (CI 0.44-1.41)</td>
<td>G2: 0.85 (CI 0.49-1.46)</td>
<td>G3: Reference</td>
</tr>
<tr>
<td></td>
<td>G1: 879</td>
<td>G2: 1447</td>
<td>G3: 898</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age, yrs:</td>
<td>Erosive disease, %: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.7-56.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex, % female:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Race, % white:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| <strong>Author, yr, Study Name:</strong> Bliddal et al., 2015⁷⁷ | Interventions, dose: Adherence to MTX: 32.9% took &lt;5 mg MTX per week of followup, and 43.5% took &lt;7.5 mg of MTX per week of followup. Median time from diagnosis to first MTX prescription was 0.66 (IQR, 0.26-1.80) years | Median time from diagnosis to first MTX prescription, yrs (IQR) | N/A | Overall: NR | N/A |
| | N: | 0.66 yrs (IQR, 0.26-1.80) | | | |
| | Mean age, yrs: | Baseline DAS, mean: NR | | | | |
| | 59.8 (SD, 14.4) | Baseline HAQ, median: NR | | | | |
| | | | | Overall discontinuation | | |
| | | | | After an initial loss of adherence, the remainder Danish RA patients slowly but steadily dropped out of treatment over the following years. After 10.9 years, 50 percent discontinued. | | | |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Bliddal et al., 2015</td>
<td>approx. 5.4 million inhabitants</td>
<td>Sex, % female: 72</td>
<td>MTX naive: 100</td>
<td>Discontinuation because of AEs</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Study Design: Observational</td>
<td></td>
<td>Race, % white: NR</td>
<td>Prior csDMARD use, % NR</td>
<td>Patient adherence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N: 18,703</td>
<td></td>
<td></td>
<td>Prior CS use, % 61</td>
<td>The main determinants of non-adherence were female gender, younger age, and tie from diagnosis to initiation of MTX.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration: Followed for mean of 7.8 yrs</td>
<td></td>
<td>Biologic non-responders: NR</td>
<td>Seropositive (RF or CCP) (%): NR</td>
<td>No difference in adherence to MTX was present between those managed in private practice (1,925 (IQR, 467–3,056) days) versus 1,892 (IQR, 452–3,316) days for patients treated in hospital. In those who filed more than one MTX prescription, the mean adherence time for 7.5mg MTX per week was 2,245 (IQR, 986–3,407) days.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baseline Sharp score, mean: NR</td>
<td>Erosive disease, %: NR</td>
<td>Specific AEs</td>
<td>NA (specific AEs for head-to-head trials only)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td>Interventions, dose:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boers et al., 1997;</td>
<td>Adults (aged 18 to 69 yrs</td>
<td>G1: MTX: 7.5 mg/wk with 1 mg/day folic</td>
<td>Median disease duration, mos: 4 (range: 1-24)</td>
<td>At 5 yrs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landewe et al., 2002;</td>
<td>fulfilling ACR criteria for RA with disease duration &lt; 2 yrs; active disease defined as ≥ 6 actively inflamed joints (located at ≥ 3 sites) and ≥ 2 of the following: ≥ 9 tender joints (irrespective of site), morning stiffness for ≥ 45 mins, Westernren’s ESR ≥ 28 mm in first hour; NSAID treatment for ≥3 mos; no prior use of csDMARDs (other than antimalarials) or corticosteroids</td>
<td>SSZ: 500 mg/day, increased to 2,000 mg/day over 3 wks (oral)</td>
<td>DAS28 disease activity, mean change per yr:</td>
<td>Overall AEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuyl et al., 2010;</td>
<td></td>
<td>PNL: 60 mg in wk 1, 40 mg in wk 2, 25 mg in wk 3, 20 mg in wk 4, 15 mg in wk 5, 10 mg in wk 6, 7.5 mg/wk thereafter until wk 28 when weaned off (oral)</td>
<td>G1: -0.02 (95% CI, -0.12 to 0.08)</td>
<td>G1: 72.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COBRA study</td>
<td></td>
<td>Placebo</td>
<td>G2: -0.13 (95% CI, -0.24 to -0.02)</td>
<td>G2: 62.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Setting:</td>
<td></td>
<td>G2: Placebo with 1 mg/day folic acid</td>
<td>p=0.265</td>
<td>SAEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands and Belgium, multicenter</td>
<td></td>
<td>SSZ: 500 mg/day, increased to 2,000 mg/day over 3 wks (oral)</td>
<td>Time-averaged DAS28 disease activity, mean change per yr:</td>
<td>G1: 2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design:</td>
<td></td>
<td>Placebo</td>
<td>G1: -0.07 (95% CI, -0.11 to 0.03)</td>
<td>G2: 7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td>NSAIDs and simple analgesics were allowed, but discontinuation was actively pursued; ≤ 2 intra-articular steroid injections were allowed in 2 periods after wk 38 (not in 6-wk period preceding independent assessment); any other intervention with parenteral or oral corticosteroids was not allowed</td>
<td>G2: -0.17 (95% CI, -0.23 to 0.11)</td>
<td>Overall discontinuation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td>155</td>
<td>MTX naive, %: 100</td>
<td>p=0.014</td>
<td>G1: 8.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>5 yrs</td>
<td>MTX inadequate responders, %: 0</td>
<td>ACR response, %: NR</td>
<td>G2: 29.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biologic non-responders, %: NA</td>
<td>DAS remission, %: NR</td>
<td>p=0.0008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RF seropositive, %: 74.4</td>
<td>Sharp score, mean change per yr:</td>
<td>Discontinuation due to AEs:</td>
<td>G1: 2.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 5.6 (95% CI, 4.3 to 7.1)</td>
<td>G2: 7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 8.6 (95% CI, 6.2 to 11)</td>
<td>p=0.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean change per yr:</td>
<td>Discontinuation due to lack of efficacy:</td>
<td>G1: 1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 0.01 (95% CI -0.03 to 0.05)</td>
<td>G2: 15.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 0.01 (95% CI -0.03 to 0.05)</td>
<td>p=0.875</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36:</td>
<td>Patient adherence (satisfactory compliance):</td>
<td>G1: 84.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td>G2: 84.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>At 80 wks</td>
<td>Specific AEs:</td>
<td>G1: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sharp score, median: G1: 4 (range: 0-80)</td>
<td>Rash:</td>
<td>G2: 5.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 12 (range: 0-72)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.01</td>
<td>GI complaints:</td>
<td>G1: 14.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 12.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr: Boers et al., 1997; Landewe et al., 2002; Tuyl et al., 2010</td>
<td>COBRA study (continued)</td>
<td>Mean age, yrs: G1: 49.5 (SD, 11.9) G2: 49.4 (SD, 12.3) Overall: NR</td>
<td></td>
<td>At 56 wks DAS28 disease activity, mean change: G1: -1.4 (SD, 1.2) G2: -1.3 (SD, 1.4) p=0.78</td>
<td>Dyspnea (final diagnosis exacerbation of chronic bronchitis): G1: 1.3 G2: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female: 58.3</td>
<td></td>
<td>Pooled index, mean change: G1: 1.1 (SD, 0.8) G2: 0.9 (SD, 0.8) p=0.20</td>
<td>Thrombocytopenia (diagnosis preleukaemic disease): G1: NR G2: 1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: 98.7</td>
<td></td>
<td>Persisting ACR-defined remission, %: G1: 1.3 G2: 4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: 0.0</td>
<td></td>
<td>Sharp score, median: G1: 2 (range: 0-43) G2: 6 (range 0-54) p=0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino: 0.0</td>
<td></td>
<td>HAQ, mean change: G1: -0.8 (SD, 0.8) G2: -0.6 (SD, 0.7) p&lt;0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pain (visual analogue scale), mean change: G1: -23 (SD, 29) G2: -25 (SD, 28) p&lt;0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boers et al., 1997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landewe et al., 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuyl et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COBRA study (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At 28 wks
DAS28 disease activity, mean change:
G1: -2.1 (SD, 1.2)
G2: -1.3 (SD, 1.2)
p<0.0001

Pooled index, mean change:
G1: 1.4 (SD, 0.7)
G2: 0.8 (SD, 0.7)
(p<0.0001)

ACR20 response, %:
G1: 72.4
G2: 49.4
p=0.006

ACR50 response, %:
G1: 48.7
G2: 26.6
p=0.007

ACR-defined probable or definite remission, %:
G1: 27.6
G2: 16.5
p=0.14

Sharp score, median:
G1: 1 (range: 0-28)
G2: 4 (range: 0-44)
p<0.0001

HAQ, mean change:
G1: -1.1 (SD, 0.8)
G2: -0.6 (SD, 0.6)
p<0.0001
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr:</td>
<td>Boers et al., 1997 molded; Landewe et al., 2002; Tuyl et al., 2010</td>
<td>COBRA study (continued)</td>
<td></td>
<td>Pain (visual analogue scale), mean change: G1: -34 (SD, 25) G2: -20 (SD, 30) p&lt;0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p<0.002
### Study Characteristics

| Author, yr, Study Name: Breedveld et al., 2006; Kimel et al., 2008; van Vollenhoven et al., 2010; van der Heijde et al., 2010; Strand et al., 2012; Smolen et al., 2013; Keystone et al., 2014; Landewe et al., 2015 PREMIER |
| Study Population Summary: Adults (aged ≥ 18 yrs) fulfilling ACR criteria for RA with disease duration <3 yrs, ≥28 swollen joints, ≥10 tender joints, ESR ≥28 mm/hr or CRP ≥1.5 mg/dl, and either rheumatoid factor positivity or ≥1 joint erosion; patients were MTX, cyclophosphamide, cyclosporine and azathioprine naïve, but could have prior treatment with ≤2 other DMARDs |

### Interventions and Patient Characteristics

| Interventions, dose: |
| G1: MTX: Initiated at 7.5 mg/wk, increased to 15 mg/wk for wks 4-8, and increased to 20 mg/wk at wk 9 (oral) |
| ADA: 40 mg every other wk (subcutaneous) |
| Folic acid: 5-10 mg/wk |
| Placebo |
| G2: ADA: 40 mg every other wk (subcutaneous) |
| Folic acid: 5-10 mg/wk |
| Placebo |
| G3: MTX: Initiated at 7.5 mg/wk, increased to 15 mg/wk for wks 4-8, and increased to 20 mg/wk at wk 9 (oral) |
| Placebo |
| Folic acid: 5-10 mg/wk |

For patients who did not achieve ACR20 response at wk 16, the injectable medication (ADA or placebo) was increased to weekly after the oral medication (MTX or placebo) was optimized.

### Baseline Disease and Treatment Characteristics

| Mean disease duration, yrs: 0.7-0.8 |
| Baseline DAS28, mean: 6.3-6.4 |
| Baseline HAQ-DI, mean: 1.5-1.6 |
| Prior csDMARD use, %: 32.4 |
| MTX naïve, %: 100 |
| MTX inadequate responders, %: 0 |

### Health Outcomes

| At 2 yrs |
| DAS28 (CRP) disease activity, mean change: |
| G1: -3.8 |
| G2: -3.1 |
| G3: -3.1 |
| ACR20 response, %: |
| G1: 69 |
| G2: 49 |
| G3: 56 |
| G1 vs. G2: p<0.001 |
| G1 vs. G3: p=0.002 |
| ACR50 response, %: |
| G1: 59 |
| G2: 37 |
| G3: 43 |
| G1 vs. G2/3: p<0.001 |
| ACR70 response, %: |
| G1: 47 |
| G2: 28 |
| G3: 28 |
| G1 vs. G2/3: p<0.001 |
| DAS28 remission (< 2.6), %: |
| G1: 49 |
| G2: 25 |
| G3: 25 |
| G1 vs. G2/3: p<0.001 |

### Adverse Events (%)

| Overall AEs: |
| G1: 97.8 |
| G2: 95.6 |
| G3: 95.3 |
| SAEs: |
| G1: 18.5 |
| G2: 21.1 |
| G3: 15.9 |
| p=0.192 |
| Overall discontinuation: |
| G1: 24.3 |
| G2: 39.1 |
| G3: 34.2 |
| p<0.001 |

### Discontinuation because of AEs:

| G1: 11.9 |
| G2: 9.5 |
| G3: 7.4 |
| p=0.21 |

### Discontinuation because of lack of efficacy:

| G1: 4.9 |
| G2: 19.0 |
| G3: 17.9 |

### Patient adherence:

| NR |

### ROB Rating

| Medium |

---

**Note:**

- **Study Name:** The study names include multiple authors and years, indicating a collaborative or multi-center study approach.
- **Interventions and Patient Characteristics:** The interventions include MTX, ADA, and folic acid, with doses and treatment regimens specified.
- **Baseline Disease and Treatment Characteristics:** Baseline characteristics such as disease duration, mean DAS28, and HAQ-DI scores are provided.
- **Health Outcomes:** Outcomes include disease activity change, ACR20, 50, and 70 responses, and DAS28 remission.
- **Adverse Events (%):** The table lists overall AEs, SAEs, and overall discontinuation rates, with p-values indicating statistical significance.
- **Patient adherence:** This is marked as NR, indicating no reported adherence data.
- **ROB Rating:** The rating is marked as Medium, suggesting a moderate level of risk of bias in the study.
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr,</td>
<td>BREEDVELD ET AL., 2006;</td>
<td>N:</td>
<td>Biologic non-responders, %:</td>
<td>Modified Sharp score</td>
<td>Specific AEs</td>
<td></td>
</tr>
<tr>
<td>Study Name:</td>
<td>KIMEL ET AL., 2008;</td>
<td>G1: 268</td>
<td>NR</td>
<td>Mean change:</td>
<td>Infections, n (per 100 patient yrs):</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VAN VOLLENHOVEN ET AL.,</td>
<td>G2: 274</td>
<td>Prior CS use, %:</td>
<td>G1: 1.9</td>
<td>G1: 123</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2010; van der HEIJDE</td>
<td>G3: 257</td>
<td>35.9</td>
<td>G2: 5.5</td>
<td>G2: 110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ET AL., 2010; STRAND ET</td>
<td>Mean age, yrs:</td>
<td>RF or CCP seropositive, %:</td>
<td>G3: 10.4</td>
<td>G3: 119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AL., 2012; SMOLEN ET AL.</td>
<td>G1: 51.9 (SD, 14.0)</td>
<td>NR</td>
<td>G1 vs. G2/3: p&lt;0.001</td>
<td>G1: 2.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2013; KEYSTONE ET AL.,</td>
<td>G2: 52.1 (SD, 13.5)</td>
<td>No radiographic progression (change ≤ 0.5), %:</td>
<td>G2 vs. G3: p&lt;0.001</td>
<td>G2: 0.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2014; LANDEWE ET AL., 2015</td>
<td>G3: 52.0 (SD, 13.1)</td>
<td></td>
<td>G1 vs. G2/3: p&lt;0.01</td>
<td>G3: 1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PREMIER (continued)</td>
<td>Overall: NR</td>
<td></td>
<td>G2 vs. G3: p&lt;0.01</td>
<td>G1 vs. G2: p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td>799</td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3: Not significant</td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>2 yrs</td>
<td></td>
<td></td>
<td></td>
<td>Malignancies, n (per 100 patient yrs):</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.9</td>
<td></td>
</tr>
<tr>
<td>Baseline mTSS, mean:</td>
<td>18.1-21.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erosive disease, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI, mean change:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI response (change ≥ 0.22), %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI score of 0 (no functional impairment), %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breedveld et al., 2006;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimel et al., 2008;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Vollenhoven et al., 2010;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Heijde et al., 2010;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strand et al., 2012;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al., 2013;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keystone et al., 2014;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landewe et al., 2015;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREMIEER (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study Population Summary</th>
</tr>
</thead>
</table>

| Interventions and Patient Characteristics |

<table>
<thead>
<tr>
<th>Baseline Disease and Treatment Characteristics</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SF-36</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental component, mean:</td>
</tr>
<tr>
<td>G1: 51.8 (SD, 8.8)</td>
</tr>
<tr>
<td>G2: 49.8 (SD, 8.1)</td>
</tr>
<tr>
<td>G3: 52.4 (SD, 8.4)</td>
</tr>
<tr>
<td>G1 vs. G3: p=0.7609</td>
</tr>
<tr>
<td>G2 vs. G3: p=0.0148</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical component, mean:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 48.8 (SD, 8.3)</td>
</tr>
<tr>
<td>G2: 44.7 (SD, 8.0)</td>
</tr>
<tr>
<td>G3: 45.9 (SD, 7.8)</td>
</tr>
<tr>
<td>G1 vs. G3: p&lt;0.0001</td>
</tr>
<tr>
<td>G2 vs. G3: p=0.3912</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pain (visual analog scale), mean:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 9.6 (SD, 14.9)</td>
</tr>
<tr>
<td>G2: 19.6 (SD, 16.5)</td>
</tr>
<tr>
<td>G3: 12.5 (SD, 15.8)</td>
</tr>
<tr>
<td>G1 vs. G2: p&lt;0.0001</td>
</tr>
<tr>
<td>G2 vs. G3: p=0.1571</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retained or gained employment, %:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 57.6 (of 210)</td>
</tr>
<tr>
<td>G3: 47.6 (of 210)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Missed work days, mean:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 17.4 (for 130 employed)</td>
</tr>
<tr>
<td>G3: 36.9 (for 110 employed)</td>
</tr>
<tr>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Author, yr, Study Name: Breedveld et al., 2006; Kimel et al., 2008; van Vollenhoven et al., 2010; van der Heijde et al., 2010; Strand et al., 2012; Smolen et al., 2013; Keystone et al., 2014; Landewe et al., 2015 PREMIER (continued)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Interventions and Patient Characteristics</td>
</tr>
<tr>
<td>Baseline Disease and Treatment Characteristics</td>
</tr>
<tr>
<td>Health Outcomes</td>
</tr>
<tr>
<td>Adverse Events (%)</td>
</tr>
<tr>
<td>ROB Rating</td>
</tr>
<tr>
<td>At 76 wks SF-36</td>
</tr>
<tr>
<td>Mental component, mean:</td>
</tr>
<tr>
<td>G1: 51.4 (SD, 8.7)</td>
</tr>
<tr>
<td>G2: 49.3 (SD, 8.1)</td>
</tr>
<tr>
<td>G3: 51.7 (SD, 8.4)</td>
</tr>
<tr>
<td>Physical component, mean:</td>
</tr>
<tr>
<td>G1: 47.5 (SD, 8.8)</td>
</tr>
<tr>
<td>G2: 43.9 (SD, 7.8)</td>
</tr>
<tr>
<td>G3: 44.7 (SD, 8.0)</td>
</tr>
<tr>
<td>Pain (visual analog scale), mean:</td>
</tr>
<tr>
<td>G1: 13.1 (SD, 15.0)</td>
</tr>
<tr>
<td>G2: 22.2 (SD, 16.9)</td>
</tr>
<tr>
<td>G3: 16.4 (SD, 16.1)</td>
</tr>
<tr>
<td>At 1 yr DAS28 (CRP) disease activity, mean change:</td>
</tr>
<tr>
<td>G1: -3.6</td>
</tr>
<tr>
<td>G2: -2.8</td>
</tr>
<tr>
<td>G4: -2.8</td>
</tr>
<tr>
<td>ACR20 response, %:</td>
</tr>
<tr>
<td>G1: 73</td>
</tr>
<tr>
<td>G2: 54</td>
</tr>
<tr>
<td>G3: 63</td>
</tr>
<tr>
<td>G1 vs. G2: p&lt;0.001</td>
</tr>
<tr>
<td>G1 vs. G3: p=0.022</td>
</tr>
<tr>
<td>ACR50 response, %:</td>
</tr>
<tr>
<td>G1: 62</td>
</tr>
<tr>
<td>G2: 41</td>
</tr>
<tr>
<td>G3: 46</td>
</tr>
<tr>
<td>G1 vs. G2/3: p&lt;0.001</td>
</tr>
<tr>
<td>Study Characteristics</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td><strong>Author, yr,</strong></td>
</tr>
<tr>
<td>Breedveld et al.,</td>
</tr>
<tr>
<td>2006&lt;sup&gt;15&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Kimel et al.,</td>
</tr>
<tr>
<td>2008&lt;sup&gt;10&lt;/sup&gt;;</td>
</tr>
<tr>
<td>van Vollenhoven et</td>
</tr>
<tr>
<td>al., 2010&lt;sup&gt;19&lt;/sup&gt;;</td>
</tr>
<tr>
<td>van der Heijde et al., 2010&lt;sup&gt;13&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Strand et al.,</td>
</tr>
<tr>
<td>2012&lt;sup&gt;116&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Smolen et al.,</td>
</tr>
<tr>
<td>2013&lt;sup&gt;117&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Keystone et al.,</td>
</tr>
<tr>
<td>2014&lt;sup&gt;118&lt;/sup&gt;;</td>
</tr>
<tr>
<td>Landewe et al.,</td>
</tr>
<tr>
<td>2015&lt;sup&gt;119&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>ACR70 response, %:</strong></td>
</tr>
<tr>
<td>G1: 46</td>
</tr>
<tr>
<td>G2: 26</td>
</tr>
<tr>
<td>G3: 28</td>
</tr>
<tr>
<td>G1 vs. G2/3: p&lt;0.001</td>
</tr>
<tr>
<td><strong>DAS28 remission (&lt; 2.6), %:</strong></td>
</tr>
<tr>
<td>G1: 43</td>
</tr>
<tr>
<td>G2: 23</td>
</tr>
<tr>
<td>G3: 21</td>
</tr>
<tr>
<td>G1 vs. G2/3: p&lt;0.001</td>
</tr>
<tr>
<td><strong>Modified Sharp score</strong></td>
</tr>
<tr>
<td>Mean change:</td>
</tr>
<tr>
<td>G1: 1.3</td>
</tr>
<tr>
<td>G2: 3.0</td>
</tr>
<tr>
<td>G3: 5.7</td>
</tr>
<tr>
<td>G1 vs. G2: p=0.002</td>
</tr>
<tr>
<td>G1 vs. G3: p&lt;0.001</td>
</tr>
<tr>
<td>G2 vs. G3: p&lt;0.001</td>
</tr>
<tr>
<td><strong>No radiographic progression</strong></td>
</tr>
<tr>
<td>(change ≤ 0.5), %:**</td>
</tr>
<tr>
<td>G1: 64</td>
</tr>
<tr>
<td>G2: 51</td>
</tr>
<tr>
<td>G3: 37</td>
</tr>
<tr>
<td>G1 vs. G2/3: p&lt;0.01</td>
</tr>
<tr>
<td>G2 vs. G3: p&lt;0.01</td>
</tr>
<tr>
<td><strong>HAQ-DI, mean change:</strong></td>
</tr>
<tr>
<td>G1: -1.1 (SD, 0.6)</td>
</tr>
<tr>
<td>G2: -0.8 (SD, 0.7)</td>
</tr>
<tr>
<td>G3: -0.8 (SD, 0.7)</td>
</tr>
<tr>
<td>G1 vs. G2, p=0.002</td>
</tr>
<tr>
<td>G1 vs. G3: p&lt;0.001</td>
</tr>
<tr>
<td>Study Characteristics</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Breedveld et al., 2006; Kimel et al., 2008; van Vollenhoven et al., 2010; van der Heijde et al., 2010; Strand et al., 2012; Smolen et al., 2013; Keystone et al., 2014; Landewe et al., 2015</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Breedveld et al., 2006; Kimel et al., 2008; van Vollenhoven et al., 2010; van der Heijde et al., 2010; Strand et al., 2012; Smolen et al., 2013; Keystone et al., 2014; Landewe et al., 2015</td>
</tr>
<tr>
<td>At 3 mos SF-36 Mental component, mean: G1: 49.7 (SD, 8.7) G2: 47.9 (SD, 8.2) G3: 550.1 (SD, 8.8)</td>
</tr>
<tr>
<td>Physical component, mean: G1: 44.8 (SD, 8.0) G2: 39.9 (SD, 7.8) G3: 41.0 (SD, 8.1)</td>
</tr>
<tr>
<td>Pain (visual analog scale), mean: G1: 23.2 (SD, 16.5) G2: 34.2 (SD, 17.9) G3: 33.8 (SD, 17.9)</td>
</tr>
<tr>
<td>Study Characteristics</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
</tr>
<tr>
<td>Study Design:</td>
</tr>
<tr>
<td>Overall N</td>
</tr>
<tr>
<td>Duration of study</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

HAQ NR
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Burmester et al., 2016-7^12,134</td>
<td>FUNCTION (continued)</td>
<td></td>
<td>SF-36</td>
<td>At 1 yr</td>
<td>DAS28-ESR, LDA, %</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td>G1: 47.6</td>
<td>G2: 57.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 50.3</td>
<td>G4: 30.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR20 response, %</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 65.3</td>
<td>G2: 67.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 65.4</td>
<td>G4: 58.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR50 response, %</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 54.9</td>
<td>G2: 56.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 50.7</td>
<td>G4: 41.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR70 response, %</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 37.8</td>
<td>G2: 43.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 37.0</td>
<td>G4: 29.3</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Burmester et al., 2016-7</td>
<td>Overall N: 1162</td>
<td>N: G1: 290, G2: 291, G3: 292, G4: 289</td>
<td>MTX inadequate responders: 0</td>
<td>DAS remission, % (95% CI)</td>
<td>Overall discontinuation at 52 wks</td>
<td></td>
</tr>
<tr>
<td>FUNCTION (continued)</td>
<td>Study Duration: 2 yrs</td>
<td>Mean age, yrs: 49.5-51.2</td>
<td>Biologic non-responders: 0</td>
<td>G1: 34.0 (28.6 to 39.5)</td>
<td>G1: 20.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female: 78.1</td>
<td>Prior CS use, %: NR</td>
<td>G2: 49.0 (43.2 to 54.7)</td>
<td>G2: 22.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td>RF seropositive, %: 89.5</td>
<td>G3: 39.4 (33.8 to 45.0)</td>
<td>G3: 19.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baseline Sharp score, mean: 5.66-7.72</td>
<td>Erosive disease, %: 100</td>
<td>G4: 19.5 (14.9 to 24.1)</td>
<td>G4: 21.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in total mTSS, mean (SD)</td>
<td>Discontinuation because of AEs at 52 wks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.42 (2.93)</td>
<td>G1: 12.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.08 (2.09)</td>
<td>G2: 20.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.26 (1.88)</td>
<td>G3: 11.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 1.14 (4.03)</td>
<td>G4: 7.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G4: p=0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI, mean change from baseline</td>
<td>Patient adherence</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.75</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.81</td>
<td>Specific AEs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: -0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: -0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G4: p=0.0024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 change from baseline</td>
<td></td>
<td></td>
<td></td>
<td>Significantly greater change in SF-36 PCS scores in the TCZ 8 mg/kg + MTX group than in the MTX group (p=0.0066).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure only; G2 &gt; G4: p=0.0066</td>
<td></td>
<td></td>
<td></td>
<td>No differences in SF-36 PCS scores between the TCZ 4 mg/kg + MTX group and the MTX group or between TCZ and MTX group.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No differences in SF-36 MCS scores.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Burmester et al., 2016-7</td>
<td>32, 134</td>
<td>FUNCTION (continued)</td>
<td>At 24 weeks DAS disease activity Figure only ACR20 response, % Figure only ACR50 response, % Figure only ACR70 response, % Figure only</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Burmester et al., 2016-7</td>
<td>FUNCTION (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DAS remission, %
- G1: 31.9
- G2: 44.8
- G3: 38.7
- G4: 15.0
  p=0.0001

Change in modified total score, mean (SD)
Figure only

HAQ-DI change from baseline
Figure only

SF-36 change from baseline
Figure only;
- G2 > G4: p=0.0014

Significantly greater change in SF-36 PCS scores in the TCZ 8 mg/kg + MTX group than in the MTX group (p=0.0014).

No differences in SF-36 PCS scores between the TCZ 4 mg/kg + MTX group and the MTX group or between TCZ and MTX group.

No differences in SF-36 MCS scores.
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Choy et al., 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> England/Wales, outpatient clinics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study design</strong> RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N</strong> 232 (out of 467 total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Duration of study</strong> 2 yrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults (aged ≥ 18 yrs) with active RA as determined by ACR criteria of &lt;24 mos and three of the following: ≥3 swollen joints, ≥ 6 tender joints, ≥45 min morning stiffness, ESR ≥28 mm/h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Interventions, dose: G1:  
  - MTX: 7.5 mg/wk, increasing incrementally to target dose of 15 mg/wk (open-label)  
  - Ciclosporin placebo  
  - PNL placebo  
G2:  
  - MTX: 7.5 mg/wk, increasing incrementally to target dose of 15 mg/wk (open-label)  
  - Ciclosporin placebo  
  - PNL: step-down initiated with MTX, initial dose of 60 mg/day and reduced to 7.5 mg at wk 6, 7.5 mg/day from wks 6 to 28, stopped by wk 34 |
| Mean disease duration, mos: 2.7-5.1 |
| Baseline DAS28, mean: 5.6-5.9 |
| Baseline HAQ, mean: 1.5-1.7 |
| Mean disease activity, mean change:  
  G1: -1.42 (SE 0.17)  
  G2: -1.37 (SE 0.15)  
ACR response, %: NR  
DAS28 remission (< 2.6), %:  
  G1: 17.9  
  G2: 20.0  
Sharp score: NR |
| **At 2 yrs** |
| **Overall AEs:**  
  SAEs:  
  G1: 17.9  
  G2: 16.5  
Overall discontinuation:  
  G1: 16.2  
  G2: 47.0  
  NNH for any adverse event leading to discontinuation was and 14 (95% CI, 6 to 65) with added PNL |
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due exclusively to toxicity:</td>
<td>G1: 6.8</td>
</tr>
<tr>
<td>Study Name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 12.2</td>
<td></td>
</tr>
<tr>
<td>Choy et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Toxicity implicated in discontinuation:</td>
<td>G1: 10.3</td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 19.1</td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due exclusively to lack of efficacy:</td>
<td>G1: 11.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 8.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lack of efficacy implicated in discontinuation:</td>
<td>G1: 16.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 11.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Patient adherence:</td>
<td>NR</td>
</tr>
</tbody>
</table>

N: G1: 117, G2: 115
Mean age, yrs: 54
Sex, % female: 69.6
Race, % white: NR

Prior CS use, %: NR
MTX naive: NR
Prior csDMARD use, %: 13.9
MTX inadequate responders: NR
Biologic non-responders: NR
RF seropositive, %: 66.8
Baseline Sharp score, mean: NR
Baseline Larsen score, median: G1: 7 (IQR, 3, 15), G2: 6 (IQR, 2, 20)
Erosive damage, %: 33.0

Cases with new erosions (primary outcome), %: G1: 29, G2: 16
Larsen score, mean change: G1: 7.41 (SE 0.99), G2: 4.70 (SE 0.69)
HAQ, mean change: G1: -0.29 (SE 0.07), G2: -0.28 (SE 0.07)
SF-36 Physical component, mean change: G1: 5.8 (SE 1.0), G2: 3.5 (SE 1.0)
Mental component: “No differences”
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Choy et al., 2008</td>
<td>Race, % black NR</td>
<td></td>
<td></td>
<td>EuroQol:</td>
<td>Specific AEs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnicity, % Latino NR</td>
<td></td>
<td></td>
<td>“No differences” at 6 mos</td>
<td>Death:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 disease activity, mean change:</td>
<td>G1: 0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -1.14</td>
<td>G2: 0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -1.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission (&lt; 2.6), %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.21</td>
<td>G2: 1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Myocardial infarctions, angina, strokes:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.9</td>
<td>G2: 1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Upper GI:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.9</td>
<td>G2: 2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Infections</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 6.0</td>
<td>G2: 3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Transient creatine elevation:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 4.3</td>
<td>G2: 3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tuberculosis:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall: 0.4</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Choy et al., 2008 (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Pneumonia:**
  - Overall: 0.6
- **Respiratory tract infection:**
  - G1: 46.1
  - G2: 42.6
- **Nausea or vomiting:**
  - G1: 12.8
  - G2: 17.4
- **Abdominal pain:**
  - G1: 6.0
  - G2: 7.8
- **Mouth ulcer:**
  - G1: 4.3
  - G2: 3.5
- **Headache:**
  - G1: 5.1
  - G2: 8.7
- **Dizziness:**
  - G1: 3.4
  - G2: 5.2
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
</table>
| Author, yr, Study Name: Choy et al., 2008 | Patients aged ≥18 with active RA (defined as >6 tender/painful joints and >6 swollen joints), ≤2 years duration, ESR >28 mm/h, or CRP >7 mg/L. | Interventions, dose:  
G1: Tofacitinib 20 mg/d + MTX (starting at 10 mg/wk, max 20 mg/wk)  
G2: Tofacitinib 20 mg/d + placebo  
G3: MTX (starting at 10 mg/wk, max 20 mg/wk) + placebo  
Tofacitinib: Administered orally as 2 5mg capsules, twice daily | Mean disease duration, yrs: 0.6-0.8  
Baseline DAS, mean: 6.3-6.5  
Baseline HAQ, mean: 1.5 | At 12 months  
DAS28-4(ESR) ≤3.2 disease activity  
G1: 58.8 (SE 8.4, p<0.001)  
G2: 30.6 (SE 7.7)  
G3: 18.9 (SE 6.4)  
ACR20 response, %  
G1: 82.9 (SE 6.4)  
G2: 66.7 (SE 7.9)  
G3: 56.8 (SE 8.1) | Diarrhea:  
G1: 4.3  
G2: 8.7  
Paresthesia:  
G1: 2.6  
G2: 6.9  
Cough:  
G1: 6.0  
G2: 9.6  
Elevated blood pressure:  
G1: 0.8  
G2: 6.9 | Medium |
| Author, yr, Study Name: Conaghan et al., 2016 | Patients aged ≥18 with active RA (defined as >6 tender/painful joints and >6 swollen joints), ≤2 years duration, ESR >28 mm/h, or CRP >7 mg/L. | Interventions, dose:  
G1: Tofacitinib 20 mg/d + MTX (starting at 10 mg/wk, max 20 mg/wk)  
G2: Tofacitinib 20 mg/d + placebo  
G3: MTX (starting at 10 mg/wk, max 20 mg/wk) + placebo  
Tofacitinib: Administered orally as 2 5mg capsules, twice daily | Mean disease duration, yrs: 0.6-0.8  
Baseline DAS, mean: 6.3-6.5  
Baseline HAQ, mean: 1.5 | At 12 months  
DAS28-4(ESR) ≤3.2 disease activity  
G1: 58.8 (SE 8.4, p<0.001)  
G2: 30.6 (SE 7.7)  
G3: 18.9 (SE 6.4)  
ACR20 response, %  
G1: 82.9 (SE 6.4)  
G2: 66.7 (SE 7.9)  
G3: 56.8 (SE 8.1) | Diarrhea:  
G1: 4.3  
G2: 8.7  
Paresthesia:  
G1: 2.6  
G2: 6.9  
Cough:  
G1: 6.0  
G2: 9.6  
Elevated blood pressure:  
G1: 0.8  
G2: 6.9 | Medium |

C49
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Conaghan et al., 2016</td>
<td>Patients were generally MTX and biological DMARD naive</td>
<td>MTX: Starting at 10 mg/wk, to 15 mg/wk at end of month 1, and 20 mg/wk at end of month 2. Administered orally and titrated if tolerated</td>
<td>MTX naive, %: 94.5</td>
<td>ACR50 response, %&lt;br&gt;G1: 65.7 (SE 8.0, p&lt;0.01)&lt;br&gt;G2: 50.0 (SE 8.3)&lt;br&gt;G3: 35.1 (SE 7.8)</td>
<td>Overall discontinuation&lt;br&gt;G1: 22.2&lt;br&gt;G2: 25.0&lt;br&gt;G3: 43.2</td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong> 109</td>
<td></td>
<td>N: G1: 36&lt;br&gt;G2: 36&lt;br&gt;G3: 37</td>
<td>Prior CS use, % 52.3</td>
<td>ACR70 response, %&lt;br&gt;G1: 28.6 (SE 7.6)&lt;br&gt;G2: 33.3 (SE 7.9)&lt;br&gt;G3: 24.3 (SE 7.1)</td>
<td>Discontinuation because of AEs&lt;br&gt;G1: 11.1&lt;br&gt;G2: 5.6&lt;br&gt;G3: 13.5</td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 1 yr</td>
<td></td>
<td>Mean age, yrs: 47.8-50.8</td>
<td>MTX inadequate responders: 0</td>
<td>DAS28-4(ESR) &lt;2.6 remission, %&lt;br&gt;G1: 35.3 (SE 8.2, p&lt;0.05)&lt;br&gt;G2: 19.4 (SE 6.6)&lt;br&gt;G3: 13.5 (SE 5.6)</td>
<td>Patient adherence&lt;br&gt;NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female: 82.6</td>
<td>Biologic non-responders: 0</td>
<td>HAQ-DI improvement vs. baseline ≥0.22&lt;br&gt;G1: 73.5 (SE 7.6)&lt;br&gt;G2: 72.2 (SE 7.5)&lt;br&gt;G3: 73.0 (SE 7.3)</td>
<td>Rash&lt;br&gt;G1: 2.8&lt;br&gt;G2: 11.1&lt;br&gt;G3: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td>Seropositive (%): RF: 75.5</td>
<td>SF-36&lt;br&gt;NR</td>
<td>Headache&lt;br&gt;G1: 8.3&lt;br&gt;G2: 5.6&lt;br&gt;G3: 5.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Anti-CCP: 79.4</td>
<td>At 6 months&lt;br&gt;DAS28-4(ESR) ≤3.2 disease activity&lt;br&gt;G1: 41.2 (SE 8.4)&lt;br&gt;G2: 27.8 (SE 7.5)&lt;br&gt;G3: 21.6 (SE 6.8)</td>
<td>URTI&lt;br&gt;G1: 8.3&lt;br&gt;G2: 5.6&lt;br&gt;G3: 5.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline Sharp score, mean: 12.6-13.7</td>
<td></td>
<td>Bronchitis&lt;br&gt;G1: 8.3&lt;br&gt;G2: 0.0&lt;br&gt;G3: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosive disease, %: 100</td>
<td></td>
<td>Diarrhea&lt;br&gt;G1: 2.8&lt;br&gt;G2: 5.6&lt;br&gt;G3: 2.7</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Conaghan et al., 2016 (continued)</td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 77.1 (SE 7.1, P &lt;0.05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 72.2 (SE 7.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 54.1 (8.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>ACR50 response, %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 57.1 (SE 8.4, p &lt;0.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 52.8 (SE 8.3, p &lt;0.05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 27.0 (SE 7.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>ACR70 response, %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 34.3 (SE 8.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 30.6 (SE 7.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 24.3 (SE 7.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>DAS28-4(ESR) &lt;2.6, disease remission</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 29.4 (SE 7.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 13.9 (SE 5.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 13.5 (SE 5.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>SHS, modification of total score, mean change from baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 0.44 (SE 0.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: -0.14 (SE 0.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 0.93 (SE 0.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>HAQ-DI improvement vs. baseline ≥0.22</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 76.5 (SE 7.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 75.0 (SE 7.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 70.3 (SE 7.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Conaghan et al., 2016&lt;sup&gt;37&lt;/sup&gt; (continued)</td>
<td>SF-36</td>
<td>At 3 months DAS28-4(ESR) ≤3.2 disease activity G1: 32.4 (SE 8.0) G2: 30.6 (SE 7.7) G3: 16.2 (SE 6.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NR</td>
<td>ACR20 response, % G1: 77.1 (SE 7.1) G2: 66.7 (SE 7.9) G3: 56.8 (SE 8.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR50 response, % G1: 48.6 (SE 8.4) G2: 55.6 (SE 8.3, p&lt;0.05) G3: 29.7 (SE 7.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR70 response, % G1: 25.7 (SE 7.4) G2: 27.8 (SE 7.5) G3: 13.5 (SE 5.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS28-4(ESR) &lt;2.6, disease remission G1: 23.5 (SE 7.3) G2: 2.8 (SE 2.7) G3: 13.5 (SE 5.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SHS, modification of total score, mean change from baseline NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Conaghan et al., 2016 (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td>Australia, Public teaching hospital</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clinical Setting</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design</strong></td>
<td>Observational (only single arm eligible)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N</strong></td>
<td>181 (119 began triple therapy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration</strong></td>
<td>104 wks median followup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interventions, dose:
At diagnosis, patients were offered initial triple therapy:
- MTX, 10 mg/wk, up to 25 mg/wk max
- SSZ 1 g b.i.d. (up titrated over 4 wks)
- HCQ 200 mg b.i.d. (up titrated over 2 wks)

Details: According to the EAC’s response driven step-up protocol:
- Every 4 wks increase MTX by 5 mg/wk as needed. If poor response after at least 4 mo, change to LEF, MTX (15 mg/wk) and HCQ (stop SSZ).
- If poor response after taking LEF for 3 mo, apply for bDMARD if meets criteria.
- If intolerant to MTX, change MTX to LEF (in addition to SSZ and HCQ).
- If intolerant to SSZ or HCQ, use 2 drugs for 3 mo then add LEF for 3 mo.

Median disease duration, mos (IQR): 6 mos (IQR, 4-10.5)
Baseline DAS, mean: 4.62 (SD, 1.37)
Baseline HAQ, median: NR
MTX naive: NR
MTX inadequate responders: NR
Prior csDMARD use, %: NR
Prior CS use, %: NR
Biologic non-responders: NR

ROB Rating: N/A

| HAQ-DI improvement vs. baseline ≥0.22 | G1: 73.5 (SE 7.6) |
| G2: 75.0 (SE 7.2) |
| G3: 81.1 (SE 6.4) |

| SF-36 |
| NR |

| Author, yr, Study Name: | Cummins et al., 2015 |
| Country, Clinical Setting: | Australia, Public teaching hospital |
| Study Design: | Observational (only single arm eligible) |
| Overall N: | 181 (119 began triple therapy) |
| Study Duration: | 104 wks median followup |

Interventions:
At diagnosis, patients were offered initial triple therapy:
- MTX, 10 mg/wk, up to 25 mg/wk max
- SSZ 1 g b.i.d. (up titrated over 4 wks)
- HCQ 200 mg b.i.d. (up titrated over 2 wks)

Details: According to the EAC’s response driven step-up protocol:
- Every 4 wks increase MTX by 5 mg/wk as needed. If poor response after at least 4 mo, change to LEF, MTX (15 mg/wk) and HCQ (stop SSZ).
- If poor response after taking LEF for 3 mo, apply for bDMARD if meets criteria.
- If intolerant to MTX, change MTX to LEF (in addition to SSZ and HCQ).
- If intolerant to SSZ or HCQ, use 2 drugs for 3 mo then add LEF for 3 mo.

Median disease duration, mos (IQR): 6 mos (IQR, 4-10.5)
Baseline DAS, mean: 4.62 (SD, 1.37)
Baseline HAQ, median: NR
MTX naive: NR
MTX inadequate responders: NR
Prior csDMARD use, %: NR
Prior CS use, %: NR
Biologic non-responders: NR

ROB Rating: N/A

Overall:
- Of the 119 patients who commenced triple therapy, 23.5% remained on MTX, HCQ, and SSZ at last followup
- SAEs: NR
- Overall discontinuation: 76
- Discontinuation of first DMARD because of AEs: 37.8
- Patients who discontinued first DMARD due to non-adherence: 4
- Specific AEs: NA (specific AEs for head-to-head trials only)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr,</strong></td>
<td></td>
<td>If multiple intolerances to MTC, LEF, SSZ or HCQ, use AZA/CYC/Gold for 3 mo. If poor response, apply for bDMARD if meets criteria.</td>
<td><strong>RF Seropositive</strong> (%)</td>
<td>74.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Name:</strong></td>
<td></td>
<td>N: 119</td>
<td><strong>Baseline Sharp score, mean:</strong></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cummins et al.,</strong></td>
<td></td>
<td>Mean age, yrs: 52.8 (SD, 13.1)</td>
<td><strong>Erosive disease,</strong> %</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2015</strong></td>
<td></td>
<td>Sex, % female: 67.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td>Race, % white: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>de Jong et al., 2013</td>
<td>Patients age ≥18 years, arthritis ≥1 joint(s) and symptom duration &lt;1 year. Patients were not included if they were diagnosed with a crystal arthropathy, (post)infectious arthritis, or autoimmune disorder other than RA, were receiving DMARDs or corticosteroids, or had contraindications for initial study medication.</td>
<td>Interventions, dose: G1: MTX (25 mg/wk, dosage reached after 3 wks) + SSZ (2 g/d) + HCQ (400 mg/d) + GCs intramuscularly G2: MTX (25 mg/wk, dosage reached after 3 wks) + SSZ (2 g/d) + HCQ 400 mg/d) + GC oral tapering scheme G3: MTX (25 mg/wk, dosage reached after 3 wks) + GC oral tapering scheme</td>
<td>Mean disease duration, mos: 166</td>
<td>Overall, patients with ≥1 AE(s): G1: 84 G2: 88 G3: 79</td>
<td>Medium</td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td>The Netherlands, Clinical</td>
<td>MTX: Doses delivered orally GCs: Tapering scheme was 15 mg/d, wks 1-4; 10 mg/d, wks 5-6; 5 mg/d, wks 7-8; 2.5 mg/d wks 9-10) For all groups: If DAS was ≥ 2.4, medication was intensified</td>
<td></td>
<td>Baseline DAS, mean: 3.36 (SD, 0.96)</td>
<td>SAEs G1: 5 G2: 11 G3: 10</td>
<td></td>
</tr>
<tr>
<td>Study Design:</td>
<td>RCT</td>
<td>N: G1: 91 G2: 93 G3: 97</td>
<td>Baseline HAQ, mean: 1.00 (SD, 0.66)</td>
<td>ACR response, % NR</td>
<td>Overall discontinuation G1: 15 G2: 9.7 G3: 10.3</td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td>515 randomized, 281 selected</td>
<td>Mean age, yrs: 53.2</td>
<td>Baseline EQ-5D, mean 0.60 – 0.65</td>
<td>DAS sustained remission, % G1 + G2: 59 G3: 53</td>
<td>Discontinuation because of AEs G1: 1.1 G2: 0.0 G3: 2.1</td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>1 yr</td>
<td>Sex, % female: 68</td>
<td>MTX naive: NR</td>
<td></td>
<td>Patient adherence At 1 yr, 1.1% of patients were listed under “protocol violations” for “no compliance”</td>
<td></td>
</tr>
<tr>
<td>Race, % white: NR</td>
<td></td>
<td>Erosive disease, %: 17.1</td>
<td>MTX inadequate responders: NR</td>
<td></td>
<td>Headache G1: 11 G2: 14 G3: 13</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>de Jong et al., 2013⁴</td>
<td></td>
<td></td>
<td>Good EULAR response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de Jong et al., 2014¹⁴⁶</td>
<td></td>
<td></td>
<td>G1: 70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kuijper, et al., 2016¹⁴⁷</td>
<td></td>
<td></td>
<td>G2: 62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>de Jong et al., 2016¹⁴⁸</td>
<td></td>
<td></td>
<td>G3: 66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tREACH (continued)</td>
<td></td>
<td></td>
<td></td>
<td>Moderate EULAR response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No EULAR response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS&lt;1.6 remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 61</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 51</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in total mTSS, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.13 (IQR, 0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0 (IQR, 0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0 (IQR, 0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.38 (SD, 0.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.51 (SD, 0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.63 (SD, 0.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean change from baseline, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.48 (SD, -0.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.42 (SD, -0.59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: -0.47 (SD, -0.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.80 (SD 0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.79 (SD 0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.77 (SD 0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>de Jong et al., 2013†</td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuijper, et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>Paid work, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>G1: 79</td>
<td>G2: 76</td>
<td>G3: 65</td>
</tr>
<tr>
<td>tREACH (continued)</td>
<td></td>
<td></td>
<td></td>
<td>Unemployed, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2</td>
<td>G2: -8</td>
<td>G3: 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G3: p=0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Working hrs/wk, median</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 32 (IQR 4-40)</td>
<td>G2: 24 (IQR 12-40)</td>
<td>G3: 25 (IQR 4-36)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Took sick leave, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 89</td>
<td>G2: 81</td>
<td>G3: 81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Took long-term sickness, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 19</td>
<td>G2: 9</td>
<td>G3: 30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Days absent, median</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Occurrence of reduction in contract hours, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 32</td>
<td>G2: 38</td>
<td>G3: 39</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuijper, et al.,</td>
<td></td>
<td></td>
<td></td>
<td>Decrease in contract hours, median</td>
<td>G1: 18 (IQR 4-37)</td>
<td></td>
</tr>
<tr>
<td>2016;</td>
<td></td>
<td></td>
<td>At 9 mos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al.,</td>
<td></td>
<td></td>
<td>DAS disease activity, mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tREACH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Increase in contract hours, median</td>
<td>G1: 8 (IQR 4-11)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Occurrence of increase in contract hours, %</td>
<td>G1: 15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Days of lost productivity, median</td>
<td>G1: 17 (IQR 3-100)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 9 mos</td>
<td>G1: 1.50 (SD 0.77)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS disease activity, mean (SD)</td>
<td>G1: 1.63 (SD 0.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean (SD)</td>
<td>G1: 0.51 (SD 0.59)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D mean, (SD)</td>
<td>G1: 0.78 (SD 0.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2013†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2014‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuijper, et al., 2016;¹⁴⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2016¹⁴⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tREACH (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 6 mos</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity, mean (SD)</td>
<td></td>
<td>G1: 1.74 (SD 0.94)</td>
<td>G2: 1.80 (SD 0.95)</td>
<td>G3: 2.02 (SD 0.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ, mean (SD)</td>
<td></td>
<td>G1: 0.45 (SD 0.53)</td>
<td>G2: 0.53 (SD 0.56)</td>
<td>G3: 0.69 (SD 0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D mean, (SD)</td>
<td></td>
<td>G1: 0.77 (SD 0.16)</td>
<td>G2: 0.76 (SD 0.17)</td>
<td>G3: 0.74 (SD 0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 3 mos</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity, mean</td>
<td></td>
<td>G1: 1.86 (SD, 0.96)</td>
<td>G2: 1.82 (SD, 0.86)</td>
<td>G3: 2.21 (SD, 1.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity, mean change from baseline</td>
<td></td>
<td>G1: -1.39 (SD, 1.0)</td>
<td>G2: -1.54 (SD, 0.98)</td>
<td>G3: -1.19 (SD, 1.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good EULAR response, %</td>
<td></td>
<td>G1: 53</td>
<td>G2: 48</td>
<td>G3: 43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate EULAR response, %</td>
<td></td>
<td>G1: 27</td>
<td>G2: 34</td>
<td>G3: 26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>de Jong et al., 2013^4</td>
<td></td>
<td></td>
<td>No EULAR response, %</td>
<td>G1: 20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>de Jong et al., 2014^146</td>
<td></td>
<td></td>
<td></td>
<td>G2: 18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kuijper, et al., 2016^147</td>
<td></td>
<td></td>
<td></td>
<td>G3: 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>de Jong et al., 2016^148</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tREACH (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td>DAS&lt;1.6 remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td>G1: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>G2: 43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuijper, et al., 2016;147</td>
<td></td>
<td></td>
<td></td>
<td>G3: 31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Jong et al., 2016148</td>
<td></td>
<td></td>
<td></td>
<td>SHS</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>tREACH (continued)</td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.51 (SD, 0.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.52 (SD, 0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.68 (SD, 0.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean change from baseline, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.41 (SD, 0.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.40 (SD, 0.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: -0.37 (SD, 0.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.75 (SD 0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.76 (SD 0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.73 (SD 0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td>den Uyl et al., 2014; ter Wee et al., 2015; COBRA-light study</td>
<td>Adults aged ≥18 yrs, fulfilling ACR criteria for RA with disease duration &lt; 2 yrs; active disease defined as ≥ 6 swollen and tender joints and ESR of ≥28 mm/h or a global health score of ≥20 mm on a 0-100 mm VAS. All patients were glucocorticoid (GC) or DMARD naïve (other than antimalarials) More exclusion criteria: uncontrolled diabetes mellitus, heart failure, uncontrolled hypertension, ALT or AST level &gt;3x the upper limit of normal, reduced renal function,</td>
<td><strong>Interventions, dose:</strong> G1: COBRA - PNL: 60 mg/d, tapered to 7.5 mg/d in 6 wks - MTX: 7.5 mg/wk - SSZ: 1 g/d, increased to 2 g/d after 1 wk - ETN intensification required for patients who did not reach DAS &lt;1.6 at wk 26 or 39: 50 mg/wk subcutaneously G2: COBRA-Lite - PNL, 30 mg/d tapered to 7.5 mg/d in 9 wks - MTX, 10 mg/d with stepwise increments in all patients to 25 mg/wk in 9 wks - ETN intensification required for patients who did not reach DAS &lt;1.6 at wk 26 or 39: patients received ETN until wk 52 Details: Concomitant treatment with NSAIDs and intra-articular injections with GCs were permitted</td>
<td><strong>Median disease duration, mos (IQR):</strong> 16 wks (IQR: 8-30) <strong>Baseline DAS, mean:</strong> 3.95-4.13 <strong>Baseline DAS28, mean:</strong> 5.45-5.67 <strong>Baseline HAQ, mean:</strong> 1.36-1.37 <strong>MTX naïve:</strong> 100 <strong>MTX inadequate responders:</strong> NR</td>
<td><strong>At 1 yr</strong> - DAS score, mean (SD) G1: 1.70 (SD, 1.0) G2: 1.88 (SD, 1.0) B (95% CI): 0.19 (CI: -0.07 to 0.45) p=0.15 <strong>DAS28 score, mean (SD)</strong> G1: 2.49 (SD, 1.3) G2: 2.71 (SD, 1.3) B (95% CI): 0.24 (CI: -0.08 to 0.57) p=0.15</td>
<td><strong>At 52 wks</strong> Overall, ≥1 AE G1: 96 G2: 96 SAEs G1: 11.1 G2: 19.8</td>
</tr>
<tr>
<td><strong>Country:</strong></td>
<td>Netherlands multicenter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Setting:</strong></td>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong></td>
<td>164</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong></td>
<td>1 yr (and 1 yr followup)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr: den Uyl et al., 2014; ter Wee et al., 2015; COBRA-light study (continued)</td>
<td>Contraindications for GCs and a positive tuberculin skin test</td>
<td>PNL: G1: wk 1 60 mg/d wk 2 40 mg/d wk 3 30 mg/d wk 4 20 mg/d wk 5 15 mg/d wk 6 10 mg/d wk 7-28 7.5 mg/d Total: 2327.5 mg G2: wk 1 30 mg/d wk 2 20 mg/d wk 3 15 mg/d wk 4-8 10 mg/d wk 9-28 7.5 mg/d Total: 2012.5 mg</td>
<td>Anti-CCP Seropositive (%): 62-66 Baseline Sharp score, mean: 1.61-2.66 Erosive disease, %: 10-17</td>
<td>ACR Non-Responders, % G1: 23 G2: 25 OR: 1.03 (0.71 to 1.49) p=0.73 DAS clinical remission (DAS &lt;1.6), % G1: 47 G2: 38 RR: 0.85 (0.64 to 1.13) p=0.18 ACR/Boolean remission, % G1: 15 G2: 17 RR: 1.03 (0.90 to 1.18) p=0.67 Mean change in SHS G1: 0.49 (SD, 1.6) G2: 0.59 (SD, 1.4) B (95% CI): 0.18 (~0.27 to 0.63) p=0.42 HAQ, mean G1: 0.57 (SD, 0.5) G2: 0.61 (SD, 0.6) B (95% CI): 0.07 (~0.08 to 0.21) p=0.35 SF-36 NR</td>
<td>Specific AEs Leukopenia G1: 1 G2: 4 At 26 wks Overall, ≥1 AE G1: 94 G2: 90 SAEs G1: 3.7 G2: 7.4 Overall discontinuation G1: 1.2 G2: 1.2 Discontinuation due to AEs G1: 1.2 G2: 1.2 Protocol violations G1: 24 G2: 7 Major protocol violations G1: 7.4 G2: 2.5 Specific AEs AEs are listed, but categories are too broad to determine specifics (i.e. “skin problems” but not rash)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td><strong>ETN:</strong></td>
<td></td>
<td><strong>At 26 wks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>den Uyl et al., 2014;</td>
<td></td>
<td>In both groups, ETN use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ter Wee et al., 2015;</td>
<td></td>
<td>stopped at 52 wks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COBRA-light study</td>
<td></td>
<td><strong>N:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td>Baseline:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Followup:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean age, yrs:</strong></td>
<td></td>
<td>51-53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex, % female:</strong></td>
<td></td>
<td>67-70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race, % white:</strong></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>DAS score, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 1.62 (SD, 0.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 1.78 (SD, 1.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Change in DAS, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: -2.50 (SD, 1.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: -2.18 (SD, 1.10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1 vs. G2: 0.21 (95% CI -0.11 to 0.53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>ACR 20 response, %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>ACR50 response %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>ACR70 response, %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Good EULAR response, %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Fulfilled EULAR Non-Response Criteria, %</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>“Minimal disease activity”</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(DAS &lt;1.6), %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 41 p=NS, NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td>Remission “according to ACR/ELUAR Boolean remission criteria,” %</td>
<td>G1: 16</td>
<td>G2: 20</td>
</tr>
<tr>
<td>den Uyl et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>SHS or Larsen score</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>ter Wee et al., 2015</td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean change from baseline</td>
<td>G1: -0.8 (SD, 0.6)</td>
<td>G2: 0.8 (SD, 0.7)</td>
</tr>
<tr>
<td>COBRA-light study</td>
<td></td>
<td></td>
<td></td>
<td>95% CI (adjusted): 0.1 (CI -0.1 to 0.2)</td>
<td>p=0.49</td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detert, 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany, Private practice, hospitals, university departments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>172</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48 wks, (open label 24-48 wks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interventions, dose:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: ADA 40 mg subcutaneously every other wk for 24 wks + open label subcutaneous MTX (15 mg/wk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: Placebo subcutaneously every other wk for 24 wks + open label subcutaneous MTX (15 mg/wk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX: Administration of ADA and placebo were discontinued after wk 24, and MTX open-label monotherapy continued until wk 48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47.2-52.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median disease duration, mos:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6-1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline DAS28, mean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2-6.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline HAQ-DI score, 1-3 scale, mean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3-1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX naïve, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX inadequate responders:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seropositive (RF or CCP) (%):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Sharp score, mean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.3-11.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erosive disease, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHS erosion score, 0-280 scale, mean 2.2-4.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 48 wks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS disease activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 3.2 (SD, 1.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 3.4 (SD, 1.6, p=0.41)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 66.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 74.9 (p=0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR50 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 52.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 51.4 (p=0.88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR70 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 40.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 34.0 (p=0.40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS remission, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 42.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 36.8 (p=0.47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHS score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 6.4 (p=0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI, mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.61 (SD, 0.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0.66 (SD, 0.6, p=0.40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36, mean score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 50.0 (SD, 9.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 47.9 (SD, 9.6, p=0.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 41.4 (SD, 12.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 42.0 (SD, 10.3, p=0.79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAEs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 13.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 19.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall discontinuation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 12.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 32.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinuation because of AEs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient adherence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 87.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 67.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific AEs here</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium (DAS28, ACR response, HAQ-DI, SF-36, attrition);</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (mTSS, SHS erosion score)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Detert, 2013&lt;sup&gt;14&lt;/sup&gt; Hit HARD (continued)</td>
<td></td>
<td></td>
<td></td>
<td>At 24 wks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS disease activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.0 (SD, 1.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 3.6 (SD, 1.4, p=0.009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 79.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 67.6 (p=0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 63.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 48.7 (p=0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 48.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 26.8 (p=0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 47.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 29.5 (p=0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.49 (SD, 0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.72 (SD, 0.6, p=0.0014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36, mean score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mental</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 48.8 (SD, 9.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 48.9 (SD, 8.8, p=0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Physical</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 44.0 (SD, 11.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 39.8 (SD, 9.9, p=0.0002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dougados et al., 1999;*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maillefert et al., 2003;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Country, Setting:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland, France, Germany (France only for 5 yr), multicenter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong></td>
<td>209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong></td>
<td>5 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interventions, dose:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: SSZ + placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: MTX + placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: SSZ + MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX: 7.5 mg wkly (2.5 mg 3 times per wk). After wk 16, could be increased to 15 mg wkly if efficacy inadequate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSZ: increased to 2 grams daily by day 9. Could be increased to 3 grams daily after wk 16 of study if efficacy was inadequate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSZ + MTX: same regiments for each drug as described above</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1: 68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2: 69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3: 72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean age, yrs:</strong></td>
<td>50-52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex, % female:</strong></td>
<td>71-77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race, % white:</strong></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline Disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean disease duration:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 2.9 mos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 2.3 mos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 3.4 mos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prior csDMARD use, %:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prior CS use, %:</strong></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MTX naive, %:</strong></td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline DAS, mean:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.13-4.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline HAQ, mean:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.25-1.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MTX inadequate responders:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biologic non-responders:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Seropositive (RF or CCP) (%):</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF+: 62-75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline Sharp total damage score, mean:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.11-8.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 1 year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean DAS change:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: -1.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: -0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: -1.26 (p=0.019 from inter-group comparisons using analysis of variance)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACR20 response, %:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 65 (p=NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean change from baseline in SHS erosion score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 2.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 2.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 1.85 (p=NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean change from baseline mTSS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 4.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 4.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 3.46 (p=NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Any detectable radiological progression in SHS erosion score, %:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 7 (p=NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Any detectable radiological progression in mTSS, %:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 9 (p=NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 1 year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall discontinuation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 30.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 21.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 29.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discontinuation because of AEs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 14.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 10.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 12.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patient adherence</strong></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nausea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 49 (p=0.007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Erythema</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 4.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 0 (p=0.047)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Increased AST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3: 0 (p=0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dougados et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999;¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maillefert et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003;¹⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(continued)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At 5 years
Txt of pts with early RA with combination therapy of MTX and SSZ during first yr did not result in any long term differences in disease activity, quality of life, or structural damage compared with monotherapy with either drug used alone

**Mean DAS (SD):**
- G1 or G2: 2.2 (1.1)
- G3: 2.2 (1)
- Overall: (p=0.9)

**Mean HAQ (SD):**
- G1 or G2: 0.6 (0.6)
- G3: 0.6 (0.7)
- Overall: (p=0.9)

**Median mTSS (IQR):**
- G1 or G2: 8.5 (1.5-17.2)
- G3: 7.5 (1.1-27.3) (p=0.7)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Durez et al., 2007;18</td>
<td>Patients fulfilling ACR criteria for RA &lt; 1 yr, ≥ 6 swollen joint count, and ≥ 8 tender joint count, no prior MTX or methyl-PNL use or prior treatment with &gt;2 DMARDs, prior treatment with glucocorticoids &lt; 3 mos (and not during 1 mo prior to study)</td>
<td>Interventions, dose: G1: • IFX: 3 mg/kg at wks 0, 2, 6 and then every 8 wks until wk 46 (intravenous) • MTX: initiated 7.5 mg/wk, increased to max 20 mg/wk by wk 14 G2: • Methyl-PNL: 1 gm at wks 0, 2, 6 and then every 8 wks until wk 46 (intravenous) • MTX: initiated 7.5 mg/wk, increased to max 20 mg/wk by wk 14 G3: • MTX: initiated 7.5 mg/wk, increased to max 20 mg/wk by wk 14 IFX and methyl-PNL stopped after 1 yr as patients continued MTX treatment only</td>
<td>Mean disease duration, yrs: G1: 0.36 (SD 0.31) G2: 0.25 (SD 0.33) G3: 0.45 (SD 0.29) DAS28-CRP, mean: G1: 5.3 (SD 1.1) G2: 5.3 (SD 1.3) G3: 5.2 (SD 0.8) HAQ, mean: G1: 1.5 (SD 0.8) G2: 1.2 (SD 0.7) G3: 1.3 (SD 0.6) MTX naïve, %: 100 MTX inadequate responders, %: NA Biologic non-responders, %: NR RF seropositive, %: G1: 67 G2: 100 G3: 64 Sharp score, mean: NR Radiographic evidence of erosions, %: G1: 13 G2: 33 G3: 36</td>
<td>Overall AEs (n): G1: 15 G2: 15 G3: 19 SAEs: G1: 0.0 G2: 0.0 G3: 6.7 Overall discontinuation: G1: 6.7 G2: 6.7 G3: 14.3 Discontinuation due to AEs: G1: 6.7 G2: 0.0 G3: 0.0 Discontinuation due to lack of efficacy: G1: 0.0 G2: 6.7 G3: 0.0 Patient adherence: NR Specific AEs: Benign infection (n) G1: 80.0 (12) G2: 80.0 (12) G3: 93.3 (14) Mild hepatotoxicity (n) G1: 14.3 (2) G2: 20.0 (3) G3: 33.5 (5)</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Durez et al., 2007;¹⁸ (continued)</td>
<td>Race, % black: NR</td>
<td></td>
<td>compared with patients receiving MTX... from baseline to week 52 (P = 0.019, respectively, by Mann-Whitney U test) (Figure 4C)”</td>
<td>SF-36: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td></td>
<td>SF-36:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>At 22 wks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DAS28-CRP disease activity, mean:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 5.57 (SD 1.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2: 5.39 (SD 1.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 4.85 (SD 0.96)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No between group differences at wk 22 or another intermediate timepoint (unclear)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %: See ACR70 response below</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %: See ACR70 response below</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %: “Clinical responses assessed by the ACR 20% improvement criteria (ACR20), the ACR50, and the ACR70 at week 22 were significantly better in the IV Methyl-PNL and IFX groups compared with the MTX group (Figure 4A)”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Durez et al., 2007; 18 (continued)</td>
<td>Adults (aged ≥ 18 yrs) with diagnosis of adult-onset RA per ACR criteria; disease duration 3-24 mos, DAS28 ≥ 3.2; either Westergren ESR ≥ 28 mm/h or CRP ≥ 20 mg/L; no</td>
<td>Interventions, dose: G1: • MTX: 7.5 mg/wk (oral); dose was titrated up over 8 wks to a max of 20 mg/wk for those with tender or swollen joints • ETN: 50 mg/wk (subcutaneous) G1a: Continue MTX + ETN in yr 2 G1b: Switch to ETN only in yr 2</td>
<td>Mean disease duration, mos: 9.0 (SE 0.3)</td>
<td>HAQ, mean: <em>HAQ scores improved significantly over time in the IV Methyl-PNL (G2) and IFX group (G1) (P &lt; 0.001 by Friedman's test), with patients receiving IV Methyl-PNL experiencing significantly more improvement compared with patients receiving MTX from baseline to week 22... (P = 0.006...by Mann-Whitney U test) (Figure 4C)</em></td>
<td>Overall AEs: yr 1 G1: 89.8 G2: 89.9 yr 2 G1a: 82.0 G1b: 80.2 G2a: 78.9 G2b: 78.8 SAEs: yr 1 G1: 12.0 G2: 12.7</td>
<td>Medium</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Emery et al., 2006; Anis et al., 2009; Emery et al., 2010; Kekow et al., 2010; Dougados et al., 2014; Zhang et al., 2012</td>
<td>prior MTX, ETN, or other TNF antagonist use; and no treatment with DMARDs or corticosteroid injections 1 mo prior to baseline visit</td>
<td>G2: • MTX: 7.5 mg/wk (oral); dose was titrated up over 8 wks to a max of 20 mg/wk for those with tender or swollen joints • Placebo G2a: Switch to MTX + ETN (50 mg/wk subcutaneous) in yr 2 G2b: Continue MTX only in yr 2</td>
<td><strong>Prior csDMARD use, %:</strong> 20.8</td>
<td><strong>ACR50 response, %:</strong> G1a: 70 (p&lt;0.001 vs. G2b) G1b: 64 G2a: 66 (p=0.007 vs. G2b) G2b: 46</td>
<td>yr 2</td>
<td><strong>Overall discontinuation:</strong> yr 1 G1: 19.3 G2: 29.5</td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> Multinational</td>
<td>N: G1: 274 (a: 111, b: 111) G2: 268 (a: 90, b: 99)</td>
<td><strong>MTX naïve, %:</strong> 100</td>
<td><strong>ACR70 response, %:</strong> G1a: 57 (p&lt;0.001 vs. G2b) G1b: 44 G2a: 48 (p=0.034 vs. G2b) G2b: 32</td>
<td><strong>DAS28 remission (&lt;2.6), %:</strong> G1a: 57.4 (of 108, p=0.002 vs. G2b) G1b: 50.0 (of 108) G2a: 58.0 (of 88, p=0.003 vs. G2b) G2b: 35.1 (of 94)</td>
<td>yr 2 G1a: 6.3 G1b: 16.2 G2a: 17.8 G2b: 23.2</td>
<td>154</td>
</tr>
<tr>
<td><strong>Study design</strong> RCT</td>
<td><strong>Mean age, yrs:</strong> 51.4 (SD, 0.6)</td>
<td><strong>MTX inadequate responders:</strong> NR</td>
<td><strong>Biologic non-responders:</strong> NR</td>
<td><strong>mTSS score Change from yr 1, mean:</strong> G1a: -0.02 (95% CI, -0.32 to 0.29; p=0.006 vs. G1b) G1b: 0.11 (95% CI, -0.54 to 0.77) G2a: 0.78 (95% CI, -0.06 to 1.61) G2b: 2.07 (95% CI, 0.42 to 3.72)</td>
<td><strong>No radiographic progression (change ≤0.5), %:</strong> G1a: 89.9 (of 99, p=0.008 vs. G1b, p=0.009 vs. G2a, p&lt;0.001 vs. G2b) G1b: 74.7 (of 99) G2a: 74.7 (of 79) G2b: 67.5 (of 83)</td>
<td>yr 2 G1a: 10.2 G2: 12.7</td>
</tr>
<tr>
<td><strong>Overall N</strong> 542</td>
<td><strong>Sex, % female:</strong> 73.3</td>
<td><strong>anti-CCP seropositive, %:</strong> 66.9</td>
<td><strong>Baseline Sharp score, mean:</strong> NR</td>
<td><strong>Discontinuation due to AEs:</strong> yr 1 G1: 10.2 G2: 12.7</td>
<td>yr 2 G1a: 2.7 G1b: 4.5 G2a: 7.8 G2b: 9.1</td>
<td>109</td>
</tr>
<tr>
<td><strong>Duration of study</strong> 2 yrs</td>
<td><strong>Race, % white</strong> 87.7</td>
<td><strong>Baseline Sharp score, mean:</strong> NR</td>
<td><strong>Discontinuation due to lack of efficacy:</strong> Yr 1 G1: 3.3 G2: 9.0</td>
<td></td>
<td>Yr 2 G1a: 0.0 G1b: 6.3 G2a: 1.1 G2b: 7.1</td>
<td>156</td>
</tr>
</tbody>
</table>

COMET (continued)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>Emery et al., 2008; Anis et al., 2009; Emery et al., 2010; Kekow et al., 2010; Dougados et al., 2014; Zhang et al., 2012; COMET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI: Mean change from yr 1:</td>
<td>G1a: Not significant/NR</td>
<td>G1b: Not significant/NR</td>
<td>G2a: 0.17 (SD, 0.42, p=0.0007)</td>
<td>G2b: Not significant/NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response (≤ 0.5), %:</td>
<td>G1a: 62 (p=0.011 vs. G2b)</td>
<td>G1b: NR</td>
<td>G2a: NR</td>
<td>G2b: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At yr 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28 LDA (≤3.2), %:</td>
<td>G1: 64.2 (of 265, 95% CI, 58 to 70)</td>
<td>G2: 41.4 (of 263, 95% CI, 35 to 47)</td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS LDA (≤2.4), %:</td>
<td>G1: 73.2 (of 265, 95% CI, 67 to 79)</td>
<td>G2: 48.7 (of 263, 95% CI, 43 to 55)</td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR 20 response, %:</td>
<td>G1: 85.9 (of 256, 95% CI, 82 to 90)</td>
<td>G2: 67.1 (of 243, 95% CI, 61 to 73)</td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR50 response, %:</td>
<td>G1: 70.7 (of 256, 95% CI, 66 to 76)</td>
<td>G2: 49.0 (of 243, 95% CI, 43 to 55)</td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient adherence:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific AEs: Death</td>
<td>yr 2</td>
<td>G1: 0.4</td>
<td>G2: 0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yr 2</td>
<td>G1a: 0.0</td>
<td>G1b: 0.0</td>
<td>G2a: 0.0</td>
<td>G2b: 1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignancies</td>
<td>yr 1</td>
<td>G1: 1.5 (leukemia [1], skin cancer [3])</td>
<td>G2: 1.5 (breast cancer [3], prostate cancer [1])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yr 2</td>
<td>G1a: 0.0</td>
<td>G1b: 0.9 (basal cell cancer)</td>
<td>G2a: 5.6 (GI cancer, bladder cancer, rectal melanoma, prostate cancer, basal cell cancer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yr 2</td>
<td>G1b: 0.0</td>
<td>G1b: 0.9 (basal cell cancer)</td>
<td>G2b: 3.0 (pancreatic cancer, cancer of the chest wall and lungs, basal cell cancer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious infections:</td>
<td>yr 1</td>
<td>G1: 1.8</td>
<td>G2: 3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Emery et al., 2008; Anis et al., 2009; Emery et al., 2010; Kekow et al., 2010; Dougados et al., 2014; Zhang et al., 2012 COMET (continued)</td>
<td></td>
<td></td>
<td><strong>ACR70 response, %:</strong></td>
<td>yr 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 48.4 (of 256, 95% CI, 41 to 55)</td>
<td>G1a: 0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 28.4 (of 243, 95% CI, 22 to 34)</td>
<td>G1b: 1.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
<td>G2a: 1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>DAS28 remission (&lt; 2.6), %:</strong></td>
<td>G2b: 2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 49.8 (of 265, 95% CI, 44 to 56)</td>
<td>Cardiovascular events, n:</td>
<td>yr 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 27.8 (of 263, 95% CI, 23 to 33)</td>
<td>G1: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
<td>G2: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>DAS remission (&lt; 1.6), %:</strong></td>
<td>yr 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 51.3 (of 265, 95% CI, 45 to 57)</td>
<td>No radiographic progression (change ≤ 0.5), %:</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 27.8 (of 263, 95% CI, 23 to 33)</td>
<td>G1: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
<td>G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>mTSS score</strong></td>
<td>Respiratory events</td>
<td>yr 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change from baseline, mean:</td>
<td>Tuberculosis:</td>
<td>G1: 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.27 (95% CI, -0.13 to 0.68)</td>
<td>G2: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 2.44 (95% CI, 1.45 to 3.43)</td>
<td>Pneumonia, n:</td>
<td>G1: 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>No radiographic progression (change ≤ 0.5), %:</strong></td>
<td>G2: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 79.7 (of 246, 95% CI, 75 to 85)</td>
<td>Upper respiratory infection, n:</td>
<td>G1: 45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 58.7 (of 230, 95% CI, 53 to 65)</td>
<td>G2: 44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>HAQ, mean change:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Normal function (HAQ-DI &lt;0.5), %</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 39 (p=0.0004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>SF-36 Mental component, mean change:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 6.1 (p=NS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Emery et al., 200812; Anis et al., 2009134; Emery et al., 2010108; Kekow et al., 2010109; Dougados et al., 2014156; Zhang et al., 2012155 COMET (continued)</td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Physical component, mean change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 13.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 10.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stopped working at least once, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 8.6 (of 105)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 24.0 (of 100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absenteeism Missed workdays, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 14.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 31.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tuberculosis:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1b: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2b: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nausea or vomiting, n:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not specified, n:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Infusion/injection site reactions, n:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Demyelination or multiple sclerosis:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1b: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2b: 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Emery et al., 2008; Anis et al., 2009; Emery et al., 2010; Kekow et al., 2010; Dougados et al., 2014; Zhang et al., 2012; COMET (continued)</td>
<td></td>
<td></td>
<td></td>
<td>Worsening of RA, n:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td>G1: 2</td>
<td>G2: 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Cholelithiasis, n:</td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td>G1: 2</td>
<td>G2: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Intervertebral disc protrusion, n:</td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td>G1: 2</td>
<td>G2: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Osteoarthritis, n:</td>
<td></td>
<td></td>
<td></td>
<td>yr 1</td>
<td>G1: 0</td>
<td>G2: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Any other AEs:</td>
<td></td>
<td></td>
<td></td>
<td>Yr 1 Interstitial lung disease (2 in combined-treatment group) and hip arthroplasty (2 in MTX group).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yr 2</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Adults (aged ≥ 18 yrs) with persistent symptoms for ≤ 2 yrs, active clinical synovitis of ≥ 2 joints for ≥ 8 wks, DAS (CRP) ≥ 3.2, and anti-CCP-2 antibody positivity; patients were either MTX-naïve at study entry or had previous exposure of ≤ 10 mg/wk for ≤ 4 wks but not within 1 mo prior to enrollment</td>
<td>Interventions, dose:</td>
<td>Mean disease duration, yrs: 0.56</td>
<td>Baseline DAS28 (CRP), mean: 5.4</td>
<td>Baseline HAQ-DI, mean: 1.4</td>
<td>12month Overall AEs:</td>
</tr>
<tr>
<td>Emery et al., 2015</td>
<td></td>
<td>G1: • ABA: 125 mg/wk (subcutaneous)</td>
<td>G1: 21.8</td>
<td>G1: 21.8</td>
<td>G1: 16.0</td>
<td>G1: 16.0</td>
</tr>
<tr>
<td>AVERT</td>
<td></td>
<td>• MTX: 7.5 mg/wk, titrated to 15-20 mg/wk within 6-8 wks</td>
<td>G2: 15.5</td>
<td>G2: 16.4</td>
<td>G2: 14.7</td>
<td>G2: 12.1</td>
</tr>
<tr>
<td>Country, Clinical Setting: Multinational</td>
<td></td>
<td>• Folic acid</td>
<td>G3: 9.5</td>
<td>G3: 15.5</td>
<td>G3: 10.3</td>
<td>G3: 17.2</td>
</tr>
<tr>
<td>Study Design: RCT</td>
<td></td>
<td>G2: • ABA: 125 mg/wk (subcutaneous)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N: 351</td>
<td></td>
<td>• Folic acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration: 2 yrs</td>
<td></td>
<td>G3: • MTX: 7.5 mg/wk, titrated to 15-20 mg/wk within 6-8 wks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Folic acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean age, yrs: 47.0 (SD, 12.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean disease duration, yrs: 0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baseline DAS28 (CRP), mean: 5.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baseline HAQ-DI, mean: 1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTX naïve, %: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTX inadequate responders, %: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biologic non-responders, %: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RF seropositive, %: 95.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At 1.5 yrs (6 mos after withdrawal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS28 (CRP) disease activity: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR20 response, %: G1: 21.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR50 response, %: G1: 16.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR70 response, %: G1: 9.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS28 (CRP) remission (&lt; 2.6), %: G1: 14.8 (of 115)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discontinuation because of AEs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discontinuation because of SAEs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12month Overall AEs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 84.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 80.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 82.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAEs: G1: 6.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 12.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 7.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall discontinuation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 13.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 21.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 17.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discontinuation because of AEs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 4.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 6.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discontinuation because of SAEs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: 2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Emery et al., 2015</td>
<td>Sex, % female: 77.8</td>
<td>anti-CCP-2 positive, %: 100</td>
<td>Sharp score: NR</td>
<td>Discontinuation due to lack of efficacy: G1: 4.3, G2: 5.2, G3: 9.5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Race, % white: 84.6</td>
<td>Baseline Sharp score, mean: NR</td>
<td>HAQ-DI response (≥ 0.3), %: G1: 21.8, G2: 16.4, G3: 10.3</td>
<td>Patient adherence: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Race, % black: NR</td>
<td>Erosive disease, %: NR</td>
<td>SF-36: NR</td>
<td>Specific AEs at 12mo: Death G1: 0.8, G2: 3.4, G3: 0.0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td></td>
<td>At 1 yr (before withdrawing treatment)</td>
<td>Serous infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS28 (CRP) disease activity: Difference in change from baseline G1 vs. G3: -0.52 (95% CI, -0.74 to -0.30) G2 vs. G3: -0.26 (95% CI, -0.11 to -0.48)</td>
<td>ACR20 response, %: G1: 74.8, G2: 63.8, G3: 65.5</td>
<td>2 died during withdrawal phase in G3: uterine neoplasm, renal failure</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %: G1: 63.0, G2: 53.4, G3: 46.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Emery et al., 2015&lt;sup&gt;7&lt;/sup&gt; AVERT (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Clinical Setting: Europe, Australia, North America, and Latin America (181 sites)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design: RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong> 879&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 1 yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Interventions, dose:**
- **G1:** CZP + MTX
  - CZP: 400 mg at wks 0, 2, and 4, 200 mg every 2 wks thereafter (subcutaneous)
  - MTX: 10-25 mg/wk (increased by 5 mg every 2 wks to 25 mg or max tolerated dose by wk 8); max tolerated dose continued through wk 52 (oral)
- **G1a:** CZP + MTX patients with very early RA (< 4 mos)
- **G1b:** CZP + MTX patients with early RA (>4 mos)
- **G2:**
  - Placebo
  - MTX: 10-25 mg/wk (increased by 5 mg every 2 wks to 25 mg or max tolerated dose by wk 8); max tolerated dose

**Mean disease duration, mos:**
- **G1:** 2.9
- **G1a:** 3.0
- **G1b:** 3.0
- **G2:** 3.0

**Baseline DAS28-ESR, mean:**
- **G1:** 6.7
- **G1a:** 6.9
- **G1b:** 6.9
- **G2:** 7.0

**Moderate disease activity (DAS28-ESR >3.2 to ≤5.1), %:**
- **G1:** 96.5
- **G1a:** 96.5
- **G1b:** 96.5
- **G2:** 96.5

**High disease activity (DAS28-ESR >3.2 to ≤5.1), %:**
- **G1:** 3.5
- **G1a:** 3.5
- **G1b:** 3.5
- **G2:** 3.5

**Baseline HAQ-DI, mean:**
- **G1:** 1.6
- **G1a:** 1.6
- **G1b:** 1.6
- **G2:** 1.6

**Prior CS use, %:**
- **G1:** 32.6 (systemic)
- **G1a:** 32.6 (systemic)
- **G1b:** 32.6 (systemic)
- **G2:** 32.6 (systemic)

**ACR70 response, %:**
- **G1:** 52.1
- **G2:** 38.8
- **G3:** 34.5

**DAS28 (CRP) remission (< 2.6), %:**
- **G1:** 60.9 (of 115)
- **G2:** 42.5 (of 113)
- **G3:** 45.2 (of 115)
p=0.01 for G1>G3

**HAQ-DI response (≥ 0.3), %:**
- **G1:** 65.5
- **G2:** 52.6
- **G3:** 44.0

**At wk 52<sup>c</sup>**

**DAS28-ESR disease activity score**
- **Change from baseline, mean:**
  - **G1:** -3.6 (SE, 0.1)
  - **G2:** -3.0 (SE, 0.1)
  - **P<0.01**

**Timepoint score, mean:**
- **G1:** 3.11 (SD, 1.58)
- **G2:** 3.77 (SD, 1.68)
- **P<0.001**

**LDA (DAS28-ESR ≤3.2), %:**
- **G1:** 54.7
- **G2:** 39.4
- **P<0.001**

**ACR20 response, %:**
- **G1:** 69.0
- **G2:** 61.5
- **P=NS**

**ACR50 response, %:**
- **G1:** 61.8
- **G2:** 52.6
- **P=0.023**

**Overall AEs (≥5% in any system organ class):**
- **G1:** 79.7
- **G2:** 72.8
- **p=NS**

**SAEs:**
- **G1:** 10.6
- **G2:** 9.2
- **p=NS**

**Overall discontinuation:**
- **G1:** 24.2
- **G2:** 34.7

**Discontinuation due to AEs:**
- **G1:** 7.7
- **G2:** 7.8
- **P=NS**

**ROB Rating**
- **Medium**

**Other**
- High (KQ 2 WPS-RA work productivity outcome scores)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Emery et al., 2017&lt;sup&gt;35, 39&lt;/sup&gt; C-EARLY (continued)</td>
<td></td>
<td>continued through wk 52 (oral)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Those in either arm with DAS28-ESR &gt;3.2 at wks 20 and 24 were withdrawn to allow them to switch to a complementary medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| | | N: | G1: 660  
G2: 219 | | | |
| | | Mean age, yrs: | 50.6 | | | |
| | | Sex, % female: | 76.8 | | | |
| | | Race, % white: | NR | | | |
| | | Race, % black: | NR | | | |
| | | Ethnicity, % Latino: | NR | | | |
| | | Prior csDMARD use, %: | 0 | | | |
| | | MTX naive: | 100 | | | |
| | | MTX inadequate responders: | 0 | | | |
| | | Biologic non-responders: | 0 | | | |
| | | RF seropositive, %: | 96.8 | | | |
| | | anti-CCP seropositive (%): | 83.9 | | | |
| | | Baseline Sharp score: Median (range): | 3.0 (0 to 161); Mean: | 7.5 | | |
| | | Erosive disease, %: | 77.8 | | | |
| | | ACR70 response, %: | G1: 51.3  
G2: 39.9  
p<0.001 | | | |
| | | Sustained LDA (DAS28-ESR ≤3.2 at both wks 40 and 52), %: | G1: 43.8  
G2: 28.6  
OR (95% CI): 2.0 (1.4 to 2.8)  
p<0.001 | | | |
| | | Sustained remission (DAS28-ESR <2.6 at both wks 40 and 52), %: | G1: 28.9  
G2: 15.0  
OR (95% CI): 2.3 (1.5 to 3.5)  
p<0.001 | | | |
| | | DAS28-ESR remission (DAS28-ESR <2.6), %: | G1: 42.6  
G2: 26.8  
OR (95% CI): 2.0 (1.4 to 2.9)  
p<0.001 | | | |
| | | mTSS score Change from baseline, mean: | G1: 0.2  
G2: 1.8  
p<0.001 | | | |
| | | No radiographic progression (change from baseline mTSS ≤0.5), %: | G1: 70.3  
G2: 49.7  
OR (95% CI): 2.4 (1.7 to 3.4)  
p<0.001 | | | |
| | | Deaths resulting from AEs<sup>e,f</sup> | G1: 0.3  
G2: 0.5  
P=NR | | | |

Patient adherence: NR

Specific AEs:* Rates for most frequently reported AEs (see below) described as “similar for both treatment arms”.

Nausea: G1: 12.6  
G2: 10.1  
p=NR

URTI G1: 10.9  
G2: 5.1  
p=NR

UTI G1: 7.3  
G2: 7.4  
p=NR

Nasopharyngitis G1: 7.0  
G2: 6.0  
p=NR

Headache G1: 6.8  
G2: 3.7  
p=NR

Deaths resulting from AEs<sup>e,f</sup> G1: 0.3  
G2: 0.5  
P=NR
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>Emery et al., 2017</td>
<td>C-EARLY (continued)</td>
<td></td>
<td>HAQ-DI change from baseline, mean</td>
<td>Active tuberculosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -1.00</td>
<td>G1: 0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.82</td>
<td>G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>p=NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normative function (HAQ-DI ≤0.5) (%)</td>
<td>Latent tuberculosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 48.1</td>
<td>G1: 0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 35.7</td>
<td>G2: 0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.002</td>
<td>p=NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fatigue: BRAF-MDQ change from baseline, mean</td>
<td>Serious Infections and Infestations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -17.8 (SE 0.6)</td>
<td>G1: 3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -15.6 (SE 1.0)</td>
<td>G2: 3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=NR</td>
<td>p=NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPS-RA: Number of work days missed in last month for employed patients, mean</td>
<td>General disorders and administration site conditions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.6 (SD 2.6)</td>
<td>G1: 16.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.9 (SD 2.5)</td>
<td>G2: 12.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=NR</td>
<td>p=NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPS-RA: Number of work days with reduced productivity in last month, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1.0 (SD 3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1.8 (SD 4.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPS-RA: Interference with work productivity in last month, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1.4 (SD 2.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1.9 (SD 2.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Emery et al., 2017&lt;sup&gt;23&lt;/sup&gt;, 39 C-EARLY (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WPS-RA: Number of days with no household work in last month, mean&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1.9 (SD 5.1)</td>
<td>G2: 3.0 (SD 6.7)</td>
<td>p=NR</td>
</tr>
<tr>
<td><strong>WPS-RA: Number of days with reduced household work productivity in last month, mean&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2.1 (SD 5.3)</td>
<td>G2: 3.0 (SD 6.6)</td>
<td>p=NR</td>
</tr>
<tr>
<td><strong>WPS-RA: Number of days with hired outside help in last month, mean&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.6 (SD 3.2)</td>
<td>G2: 0.7 (SD 3.3)</td>
<td>p=NR</td>
</tr>
<tr>
<td><strong>WPS-RA: Number of days missed of family/social/leisure activities in last month, mean&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.9 (SD 3.6)</td>
<td>G2: 0.9 (SD 3.1)</td>
<td>p=NR</td>
</tr>
<tr>
<td><strong>WPS-RA: Interference with household work productivity in last month, mean&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1.9 (SD 2.5)</td>
<td>G2: 2.5 (SD 2.8)</td>
<td>p=NR</td>
</tr>
<tr>
<td><strong>At wk 40&lt;sup&gt;c&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LDA (DAS28-ESR ≤3.2), %:</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 49.2</td>
<td>G2: 32.9</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Emery et al., 2017 38, 39</td>
<td>C-EARLY (continued)</td>
<td></td>
<td>HAQ-DI change from baseline, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p≤0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>At wk 36</strong></td>
<td><strong>LDA (DAS28-ESR ≤3.2), %:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 45.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 31.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>HAQ-DI change from baseline, mean</strong></td>
<td>G1: -0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p≤0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>At wk 24</strong></td>
<td><strong>DAS28-ESR disease activity score, mean:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.54 (SD, 1.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 4.07 (SD, 1.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>LDA (DAS28-ESR ≤3.2), %:</strong></td>
<td>G1: 39.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 30.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p≤0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>HAQ-DI change from baseline, mean</strong></td>
<td>G1: -0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p≤0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>At wk 20</strong></td>
<td><strong>LDA (DAS28-ESR ≤3.2), %:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 40.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 28.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p≤0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Emery et al., 2017</td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI change from baseline, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-EARLY (continued)</td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.90</td>
<td>G2: -0.79</td>
<td>p≤0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At wk 12c</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28-ESR disease activity score, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.88 (SD, 1.44)</td>
<td>G2: 4.43 (SD, 1.46)</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LDA (DAS28-ESR ≤3.2), %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 31.6</td>
<td>G2: 18.5</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI change from baseline, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.85</td>
<td>G2: -0.69</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Note: BeST protocol uses thrice-monthly DAS calculations and aims at achieving low DAS, with a protocol that requires treatment adjustments if DAS is &gt;2.4, but stable (after 6 mos tapering off) medication as long as the DAS is ≤2.4</td>
<td>Median symptom duration, wks 23-26 wks</td>
<td>At 10 yrs Low DAS (≤2.4), %</td>
<td>At 10 yrs</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Prior CS use, % NR</td>
<td>G1: 84</td>
<td>Overall:</td>
<td>for 10 year outcomes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Prior csDMARD use, % 0</td>
<td>G2: 77</td>
<td>Overall, 89% of patients reported AEs (74 AEs per 100 patient years). These were equally distributed between the 4 groups (p=0.159) at 10 year followup</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MTX naïve, % 100</td>
<td>G3: 83</td>
<td>Patients who reported SAEs</td>
<td>Overall, 47% of patients reported SAEs (12 SAEs per 100 patient-years) at 10 year followup</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline DAS score, mean 4.3-4.5</td>
<td>G4: 84</td>
<td>SAEs per 100 patient-years, yrs 6-10</td>
<td>G1: 13.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline D-HAQ score, mean 1.4</td>
<td>ACR20/50/70 or EULAR response, % NR</td>
<td>G2: 22</td>
<td>G1: 13.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Achieved drug-free remission during at least 1 visit, %</td>
<td>G3: 22</td>
<td>G2: 10.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In drug-free remission during at 10 yrs, %</td>
<td>G4: 29</td>
<td>G3: 12.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Increase in mTSS, median (IQR)</td>
<td>G1: 2.0 (IQR, 0 to 11.0)</td>
<td>G4: 13.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mortality, at 10 yr followup</td>
<td>G2: 2.5 (IQR, 0 to 13.5)</td>
<td>CVD adverse events per 100 patient years, yrs 6-10</td>
<td>G1: 5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G3: 3.0 (IQR, 0.3 to 11.3)</td>
<td>G3: 6.4</td>
<td>G1: 12.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G4: 1.5 (IQR, 0.0 to 6.0)</td>
<td>G3: 7.8</td>
<td>G2: 12.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 5.7</td>
<td>G3: 15.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 15.6</td>
<td>G4: 15.6</td>
<td></td>
</tr>
</tbody>
</table>

**Country and setting**

- The Netherlands
- 18 peripheral and 2 university hospitals

**Study design**

RCT
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goekoop-Ruiterman, 2005&lt;sup&gt;75&lt;/sup&gt;; Allaart et al., 2006&lt;sup&gt;71&lt;/sup&gt;; Goekoop-Ruiterman, 2007&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;96&lt;/sup&gt;; Dirven et al., 2012&lt;sup&gt;80&lt;/sup&gt;; Dirven et al., 2013&lt;sup&gt;82&lt;/sup&gt;; Klarenbeek et al., 2010&lt;sup&gt;11&lt;/sup&gt;; Markusse et al., 2016&lt;sup&gt;87&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;10&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;11&lt;/sup&gt;; Markusse et al., 2014&lt;sup&gt;88&lt;/sup&gt;; BeST Study (continued)</td>
<td>Overall N: 508 Duration of study: 10 yrs</td>
<td>G6: Initial combination therapy group (iCombo): Combined G3 + G4 for post-hoc analysis G6a: poor prognosis patients from G6 G6b: non-poor prognosis patients from G6 Group details: <strong>Sequential monotherapy details:</strong> Subsequent steps for patients with insufficient response were SSZ monotherapy, LEF monotherapy, MTX with IFX, gold with methylprednisolone, and, MTX with cyclosporin A (CSA) and PRED <strong>Step-up combination therapy details:</strong> Patients whose disease failed to respond to the combination of the 4 drugs switched to MTX with IFX, MTX with CSA and PRED, and, lastly, to LEF. <strong>Initial combination therapy with PRED details:</strong> If DAS &gt;2.4, MTX increased to 25-30 mg/wk. If response still insufficient, combination replaced by combination of MTX with CSA and PRED, followed by MTX with IFX, LEF monotherapy, gold with methylprednisolone, and lastly, by azathioprine (AZA)</td>
<td>Estimated mTSS corrected for baseline, mean: G1: 11 G2: 8 G3: 8 G4: 6 p=0.15 G1 vs. G2: p=0.046 For all other comparisons: p&gt;0.10</td>
<td>Mortality, yrs 6-10: G1: 5 G2: 5 G3: 7 G4: 10</td>
<td>Mortality, after dropout: G1: 8 G2: 7 G3: 12 G4: 6</td>
<td>At 5 years Overall: G1: 87 G2: 85 G3: 84 G4: 88 p=0.84</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005; Allaart et al., 2006; Goekoop-Ruiterman, 2007; van der Kooij, 2009; Dirven et al., 2012; Dirven et al., 2013; Klarenbeek et al., 2010; Markusse et al., 2016; Klarenbeek et al., 2011; Markusse et al., 2014; BeST Study</td>
<td>with PRED. If persistent DAS of &gt;2.4, first PRED was tapered to zero after 28 weeks, then MTX tapered to zero after 40 weeks. <strong>Initial combination therapy with IFX details:</strong> After 3 mos, dose increased to 6 mg/kg/every 8 wks if DAS was &gt;2.4. If DAS was &gt;2.4, the next infusion was increased to 7.5 mg/kg/every 8 weeks and finally to 10 mg/kg/every 8 weeks. If patients still had a DAS of &gt;2.4 while receiving MTX with 10 mg/kg IFX, medication was switched to SSZ, then to LEF, then to the combination of MTX, CSA, and PRED, then to gold with methylprednisolone, and, finally, to AZA with PRED. In the case of a persistent good response (DAS of &gt;2.4 for at least 6 months), the dose of IFX was reduced (from 10 to 7.5, 6, and then 3 mg/kg) every next infusion until stopped. <strong>Number in group:</strong> G1: 126 G2: 121 G3: 133 G4: 128 Overall: 508</td>
<td>Still in DAS drug-free remission (of those who were ever in drug free remission) at yr 5, % G1: 45 G2: 58 G3: 42 G4: 58 <strong>SHS progression, median (mean):</strong> G1: 3.5 (14.0) G2: 2.3 (11.0) G3: 1.0 (7.6) G4: 1.0 (6.0) G1&amp;G2 vs. G4: P &lt;0.01 G1 vs. G3: p&lt;0.001 <strong>Changes in HAQ</strong> Figure only <strong>SF-36 Physical and Mental Component scores</strong> Figure only <strong>At 4 years</strong> LDA (DAS ≤2.4), % G1: NR G2: NR G3: NR G4: NR Overall: 81 p=0.10 <strong>ACR20/50/70 or EULAR response, %</strong> NR</td>
<td>Discontinuation because of AEs NR Patient adherence NR Specific AEs NR Vertebral Fractures In total, vertebral fractures were observed in 15% of the 275 patients who had radiographs of the spine at 5 yrs Univariate treatment variables predictive of an ALT of &gt;2x ULN, Odds ratio (95% CI) Number of DMARDs during MTX use: 0.71 (CI 0.57-0.90, p=0.005) Mean dosage of MTX over time: 1.08 (CI 1.02-1.13, p=0.003) Time on SSZ: 0.70 (CI 0.52-0.94, p=0.018) Time on IFX: 0.72 (CI 0.54-0.95, p=0.021) Time on PRED: 0.49 (CI 0.28-0.84, p=0.010) Time on HCO: 0.59 (CI 0.26-1.35, p=0.212) Time on CSA: 1.08 (CI 0.63-1.86, p=0.784)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005; Allaart et al., 2006; Goekoop-Ruiterman, 2007; van der Kooij, 2009; van der Kooij, 2009; Dirven et al., 2012; Dirven et al., 2013; Klarenbeek et al., 2010; Markusse et al., 2016; Klarenbeek et al., 2011; Klarenbeek et al., 2011; Markusse et al., 2014; BeST Study (continued)</td>
<td>Post-hoc: G5: 200 G6: 217 Post-hoc overall: 417 Mean age, years 54-55 Sex, % female 65-71 Race, % NR</td>
<td>DAS remission (&lt;1.6), % G1: 50 G2: 41 G3: 38 G4: 42 p=0.40 Drug-free remission, % G1: 14 G2: 12 G3: 8 G4: 18 Progression of SHS score from baseline, mean (SD) G1: 11.7 (SD, 17.3) G2: 9.7 (SD, 12.8) G3: 6.7 (SD, 9.6) G4: 5.4 (SD, 9.2) p=0.005 Progression of SHS score from baseline, median (IQR) G1: 5.0 (IQR, 1.0-15.8) G2: 5.5 (IQR, 1.0-13.8) G3: 3.0 (IQR, 1.0-7.5) G4: 2.5 (IQR, 0.5-6.5) p=0.005 G1 vs. G2: p=0.77; G1 vs. G3: p=0.06; G1 vs. G4: p=0.002; G2 vs. G3: p=0.10; G2 vs. G4: p=0.005; G3 vs. G4: p=0.18</td>
<td>CVD adverse events per 100 patient years, yrs 3-5 G1: 2.6 G2: 4.1 G3: 5.5 G4: 3.3 Mortality, yrs 3-5 G1: 3 G2: 2 G3: 1 G4: 2 At 4 yrs Overall (Any AE during 4 yrs) G1: 82 G2: 83 G3: 80 G4: 84 p=0.87 Overall (Any AE during yrs 3-4) G1: 60 G2: 61 G3: 57 G4: 64 p=0.72 Total SAE during 4 yrs G1: 49.6 G2: 44.6 G3: 50.4 G4: 39.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005; Allaart et al., 2006; Goekoop-Ruiterman, 2007; van der Kooij, 2009; van der Kooij, 2009; Dirven et al., 2012; Dirven et al., 2013; Klarenbeek et al., 2010; Markusse et al., 2016; Klarenbeek et al., 2011; Klarenbeek et al., 2011; Markusse et al., 2014; BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>Improvement in HAQ compared with baseline, mean (SD)</td>
<td>Overall discontinuation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.8 (SD, 0.6)</td>
<td>G1: 8.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.7 (SD, 0.8)</td>
<td>G2: 16.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.8 (SD, 0.8)</td>
<td>G3: 10.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 0.8 (SD, 0.8)</td>
<td>G4: 7.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall: 11</td>
<td>Overall: 14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 15.1</td>
<td>Cardiovascular, yrs 3-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 12.4</td>
<td>G1: 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 18.8</td>
<td>G2: 6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 8.6</td>
<td>G3: 10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall: 14</td>
<td>G4: 9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.25</td>
<td>p=0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dermal/mucosal, yrs 3-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 Physical and Mental Component scores</td>
<td>Overall discontinuation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td>G1: 8.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 16.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall: 11</td>
<td>Overall: 14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 15.1</td>
<td>Cardiovascular, yrs 3-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 12.4</td>
<td>G1: 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 18.8</td>
<td>G2: 6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 8.6</td>
<td>G3: 10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall: 14</td>
<td>G4: 9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.25</td>
<td>p=0.72</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005; Allaart et al., 2006; Goekoop-Ruiterman, 2007; van der Kooij, 2009; van der Kooij, 2009; Dirven et al., 2012; Dirven et al., 2013; Klarenbeek et al., 2010; Markusse et al., 2016; Klarenbeek et al., 2011; Klarenbeek et al., 2011; Markusse et al., 2014; BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>At 2 years LDA, DAS ≤2.4 achieved, %</td>
<td>At 2 years Overall</td>
<td>Overall, 38% of patients had at least 1 adverse event in the second year</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 75 G2: 81 G3: 78 G4: 82 Overall: 79 p=0.554</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LDA, DAS ≤2.4 at least once, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 92 G2: 97 G3: 97 G4: 93 p=0.256</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time until DAS ≤2.4, median months (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 9 (IQR, 6–12) G2: 9 (IQR, 6–12) G3: 3 (IQR, 3–6) G4: 3 (IQR, 3–6) p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Duration first DAS ≤2.4, median months (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
</tbody>
</table>
| Goekoop-Ruiterman, 2005<sup>75</sup>; Allaart et al., 2006<sup>71</sup>; Goekoop-Ruiterman, 2007<sup>83</sup>; van der Kooij, 2009<sup>83</sup>; van der Kooij, 2009<sup>85</sup>; Dirven et al., 2012<sup>100</sup>; Dirven et al., 2013<sup>82</sup>; Klarenbeek et al., 2010<sup>91</sup>; Markusse et al., 2016<sup>87</sup>; Klarenbeek et al., 2011<sup>90</sup>; Klarenbeek et al., 2011<sup>96</sup>; Markusse et al., 2014<sup>88</sup> | Goekoop-Ruiterman, 2005<sup>75</sup>; Allaart et al., 2006<sup>71</sup>; Goekoop-Ruiterman, 2007<sup>83</sup>; van der Kooij, 2009<sup>83</sup>; van der Kooij, 2009<sup>85</sup>; Dirven et al., 2012<sup>100</sup>; Dirven et al., 2013<sup>82</sup>; Klarenbeek et al., 2010<sup>91</sup>; Markusse et al., 2016<sup>87</sup>; Klarenbeek et al., 2011<sup>90</sup>; Klarenbeek et al., 2011<sup>96</sup>; Markusse et al., 2014<sup>88</sup>; BeST Study (continued) | | ACR20 response, % | At 1 year
Patients experiencing ≥1 AE
G1: 43
G2: 47
G3: 37
G4: 39
Overall: 41
p=0.367 |
<p>| | | | ACR70 response, % | | | |
| | | | DAS &lt;1.6 remission, % | | | |
| | | | G1: 46 | | | |
| | | | G2: 38 | | | |
| | | | G3: 41 | | | |
| | | | G4: 42 | | | |
| | | | DAS &lt;1.6 remission, at least once, % | | | |
| | | | G1: 70 | | | |
| | | | G2: 64 | | | |
| | | | G3: 73 | | | |
| | | | G4: 79 | | | |
| | | | p=0.374 | | | |
| | | | Time until DAS &lt;1.6, median months (IQR) | | | |
| | | | G1: 12 (IQR, 8–19) | | | |
| | | | G2: 12 (IQR, 6–18) | | | |
| | | | G3: 6 (IQR, 3–15) | | | |
| | | | G4: 6 (IQR, 6–12) | | | |
| | | | p&lt;0.001 | | | |
| | | | Duration first DAS &lt;1.6, median months (IQR) | | | |
| | | | G1: 6 (IQR, 3–15) | | | |
| | | | G2: 6 (IQR, 3–9) | | | |
| | | | G3: 6 (IQR, 3–10) | | | |
| | | | G4: 6 (IQR, 3–15) | | | |
| | | | p=0.628 | | | |
| | | | Discontinuation because of AEs | | | |
| | | | G1: 0 | | | |
| | | | G2: 0.8 | | | |
| | | | G3: 0 | | | |
| | | | G4: 0 | | | |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td>HAQ score, mean</td>
<td>Patient adherence</td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td>HAQ score, improvement from BL, mean</td>
<td>Overall, 5% discontinued adherence to protocol because of noncompliance, but not all were lost to followup, and all available data were included in the ITT analysis</td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td>Progression of SHS from baseline, mean (SD)</td>
<td>Skin rash or other mild dermal or mucosal events</td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G1: 9.0 (SD, 17.9)</td>
<td>G1: 10</td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G2: 5.2 (SD, 8.1)</td>
<td>G2: 12</td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G3: 2.6 (SD, 4.5)</td>
<td>G3: 9</td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td>G4: 2.5 (SD, 4.6)</td>
<td>G4: 6</td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td>p=0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>G1 &amp; G2 vs. G3 &amp; G4: p&lt;0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>Progression of SHS from baseline, median (IQR)</td>
<td>Infections (mainly upper respiratory tract)</td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G1: 2.0 (IQR, 0.0 - 8.6)</td>
<td>G1: 4</td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>G2: 2.0 (IQR, 0.3 - 7.0)</td>
<td>G2: 7</td>
<td></td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>G3: 1.0 (IQR, 0.0 - 2.5)</td>
<td>G3: 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 1.0 (IQR, 0.0 - 3.0)</td>
<td>G4: 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in SHS</td>
<td>Cardiovascular events in year 1, %</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td>G1: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Relative risk for SHS Progression, RR (95% CI)</td>
<td>G2: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 1.0</td>
<td>G3: 6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.91 (CI 0.73-1.12)</td>
<td>G4: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.74 (CI 0.61-0.89)</td>
<td>CVD adverse events per 100 patient years, yrs 1-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 0.73 (CI 0.61-0.88)</td>
<td>G1: 2.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 7.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 5.2</td>
<td></td>
</tr>
</tbody>
</table>

Note: ROB = Risk of Bias
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS, improvement from baseline, mean</td>
<td>Mortality, yrs 1-2</td>
<td>G1: 0</td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1</td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 2</td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0</td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1</td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1</td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 2</td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0</td>
</tr>
<tr>
<td>Markusse et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1</td>
</tr>
<tr>
<td>Markussse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1</td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS improvement from baseline, mean</td>
<td></td>
<td>G1: 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Systolic Blood pressure (mm Hg)</td>
<td></td>
<td>G1: 4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td>G2: 4.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 4.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Diastolic Blood pressure (mm Hg)</td>
<td></td>
<td>G1: 4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td>G2: 4.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 4.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 1 year</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LDA, DAS ≤2.4 reached, %</td>
<td></td>
<td>G1: 53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.004 for 1 vs. 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.001 for 1 vs. 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=NS and NR for others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td>G5a: 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td>G6a: 93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G5b: 72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G6b: 85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p=0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td>(G1-G4): NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>G5a: 57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G6a: 71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>G5b: 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>G6b: 68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p=0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(G1-G4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a: 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6a: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b: 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6b: 39</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p=0.380</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS &lt;1.6 remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(G1-G4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a: 21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6a: 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b: 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6b: 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p=1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td>Study Characteristics</td>
<td></td>
<td></td>
<td>Change in SHS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td>SHS Progression, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td>G5a: 1.5 (0 to 5.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G6a: 0 (0 to 2.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G5b: 0 (0 to 1.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G6b: 0 (0 to 1.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p=0.451</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>mTSS (0-448 scale), mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td>G1: 7.1 (SD, 15.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td>G2: 4.3 (SD, 6.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>G3: 2.0 (SD, 3.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G4: 1.3 (SD, 4.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G1 &amp; G2 vs. G3 &amp; G4: p&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td>mTSS (0-448 scale), median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study</td>
<td></td>
<td>(continued)</td>
<td></td>
<td>G1: 2.0 (IQR, 0.0-7.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 2.5 (IQR, 0.0-6.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1.0 (IQR, 0.0-2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 0.5 (IQR, 0.0-2.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 &amp; G2 vs. G3 &amp; G4: p&lt;0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ score, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only (G1-G4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only (G5a/b and G6 a/b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ score, improvement from BL, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.7 (SD, 0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.7 (SD, 0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.9 (SD, 0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 0.9 (SD, 0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3 &amp; G4: p&lt;0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005; Allaart et al., 2006; Goekoop-Ruiterman, 2007; van der Kooij, 2009; van der Kooij, 2009; Dirven et al., 2012; Dirven et al., 2013; Klarenbeek et al., 2010; Markusse et al., 2016; Klarenbeek et al., 2011; Klarenbeek et al., 2011; Markusse et al., 2014; BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>Decrease in HAQ score, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a: -0.75 (IQR, -1.13, -0.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6a: -0.88 (IQR, -1.38, -0.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p=0.110</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b: -0.83 (-1.13, -0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G6b: -0.88 (IQR, -1.25, -0.31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p&lt;0.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 PCS, improvement from baseline, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 8.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 11.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 11.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 12.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS improvement from baseline, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 4.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Systolic Blood pressure (mm Hg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Difference in systolic BP between groups (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2: 1.82 (CI –1.11 to 4.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3: 0.32 (CI –2.57 to 3.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G4: 4.83 (CI 1.98 to 7.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G3: -1.51 (CI –4.47 to 1.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G4: 3.01 (CI 0.08 to 5.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3 vs. G4: 4.51 (CI 1.67 to 7.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005&lt;sup&gt;70&lt;/sup&gt;; Allaart et al., 2006&lt;sup&gt;71&lt;/sup&gt;; Goekoop-Ruiterman, 2007&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;96&lt;/sup&gt;; Dirven et al., 2012&lt;sup&gt;160&lt;/sup&gt;; Dirven et al., 2013&lt;sup&gt;82&lt;/sup&gt;; Klarenbeek et al., 2010&lt;sup&gt;11&lt;/sup&gt;; Markusse et al., 2016&lt;sup&gt;87&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;96&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;96&lt;/sup&gt;; Markusse et al., 2014&lt;sup&gt;53&lt;/sup&gt;; BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Diastolic Blood pressure (mm Hg)**
Figure only

**Difference in diastolic BP between groups (95% CI)**
G1 vs. G2: 1.28 (CI –0.43 to 2.99)
G1 vs. G3: 2.04 (CI 0.35 to 3.73)
G1 vs. G4: 2.81 (CI 1.15 to 4.48)
G2 vs. G3: 0.76 (CI –0.97 to 2.49)
G2 vs. G4: 1.54 (CI –0.17 to 3.24)
G3 vs. G4: 0.77 (CI –0.89 to 2.44)

**At 9 months**
LDA, DAS ≤2.4 reached, %
Figure only

**ACR20 response, %**
Figure only

**ACR70 response, %**
Figure only

**DAS <1.6 remission, %**
Figure only

**Change in SHS**
Figure only

**HAQ score, mean**
Figure only (G1-G4) Figure only (G5a/b and G6 a/b)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
</table>
| Goekoop-Ruiterman, 2005<sup>76</sup>; Allaart et al., 2006<sup>71</sup>; Goekoop-Ruiterman, 2007<sup>85</sup>; van der Kooij, 2009<sup>83</sup>; van der Kooij, 2009<sup>96</sup>; Dirven et al., 2012<sup>80</sup>; Dirven et al., 2013<sup>82</sup>; Klarenbeek et al., 2010<sup>91</sup>; Markusse et al., 2016<sup>87</sup>; Klarenbeek et al., 2011<sup>90</sup>; Klarenbeek et al., 2011<sup>98</sup>; Markusse et al., 2014<sup>83</sup> | BeST Study (continued) | | HAQ score, improvement from BL, mean | G1: 0.6 (SD, 0.7)  
G2: 0.6 (SD, 0.7)  
G3: 0.8 (SD, 0.7)  
G4: 0.8 (SD, 0.6)  
p=0.010  
G1 & G2 vs. G3 & G4: p<0.050 |
<p>| SF-36 | | | | NR |
| Systolic Blood pressure (mm Hg) | | | | Figure only |
| Diastolic Blood pressure (mm Hg) | | | | Figure only |
| At 6 months | | | | LDA, DAS ≤2.4 reached, % |
| ACR20 response, % | | | | Figure only |
| ACR20 response, % | | | | Figure only |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goekoop-Ruiterman, 2005&lt;sup&gt;76&lt;/sup&gt;; Allaart et al., 2006&lt;sup&gt;71&lt;/sup&gt;; Goekoop-Ruiterman, 2007&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;96&lt;/sup&gt; Dirven et al., 2012&lt;sup&gt;80&lt;/sup&gt;; Dirven et al., 2013&lt;sup&gt;82&lt;/sup&gt;; Klarenbeek et al., 2010&lt;sup&gt;11&lt;/sup&gt;; Markusse et al., 2016&lt;sup&gt;87&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;95&lt;/sup&gt; Klarenbeek et al., 2011&lt;sup&gt;95&lt;/sup&gt; Markusse et al., 2014&lt;sup&gt;58&lt;/sup&gt; BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, % Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS &lt;1.6 remission, % Figure only</td>
<td></td>
<td></td>
<td></td>
<td>Change in SHS Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ score, mean Figure only (G1-G4) Figure only (G5a/b and G6 a/b)</td>
<td></td>
<td></td>
<td></td>
<td>HAQ score, improvement from BL, mean G1: 0.5 (SD, 0.7) G2: 0.5 (SD, 0.7) G3: 0.9 (SD, 0.7) G4: 0.8 (SD, 0.6) (p&lt;0.001) G1 &amp; G2 vs. G3 &amp; G4: p&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 PCS, improvement from baseline, mean G1: 8.0 G2: 8.5 G3: 12.5 G4: 12.4 p&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 MCS improvement from baseline, mean G1: 3.1 G2: 3.5 G3: 1.2 G4: 4.1 p=0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman,</td>
<td></td>
<td></td>
<td></td>
<td>Systolic Blood pressure (mm Hg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005&lt;sup&gt;76&lt;/sup&gt;;</td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al.,</td>
<td></td>
<td></td>
<td></td>
<td>Diastolic Blood pressure (mm Hg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006&lt;sup&gt;51&lt;/sup&gt;;</td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman,</td>
<td></td>
<td></td>
<td></td>
<td>At 3 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007&lt;sup&gt;83&lt;/sup&gt;;</td>
<td></td>
<td></td>
<td></td>
<td>LDA, DAS ≤2.4 reached, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij,</td>
<td></td>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009&lt;sup&gt;83&lt;/sup&gt;;</td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij,</td>
<td></td>
<td></td>
<td></td>
<td>Figure only (G1-G4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009&lt;sup&gt;56&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>G5a: 38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G6a: 70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012&lt;sup&gt;10&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>G5a vs. G6a: p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G5b: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013&lt;sup&gt;82&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td>G6b: 71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G5b vs. G6b: p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010&lt;sup&gt;11&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016&lt;sup&gt;87&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011&lt;sup&gt;91&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011&lt;sup&gt;99&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014&lt;sup&gt;53&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study</td>
<td></td>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allaart et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goekoop-Ruiterman, 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Kooij, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirven et al., 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klarenbeek et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Markusse et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BeST Study (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACR70 response, %
Figure only (G1-G4)
G5a: 4
G6a: 24
G5a vs. G6a: p<0.001
G5b: 3
G6b: 17
G5b vs. G6b: p=0.001

DAS <1.6 remission, %
Figure only (G1-G4)
G5a: 5
G6a: 17
G5a vs. G6a: p=0.016
G5b: 7
G6b: 18
G5b vs. G6b: p=0.017

Change in SHS
Figure only

HAQ score, mean
Figure only (G1-4)
G5: 1.08
G6: 0.60
Figure only (G5a/b and G6 a/b)

HAQ score, improvement from BL, mean
G1: 0.4 (SD, 0.6)
G2: 0.3 (SD, 0.6)
G3: 0.8 (SD, 0.7)
G4: 0.7 (SD, 0.6)
p<0.001
G1/G2 vs. G3/G4: p=0.050
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goekoop-Ruiterman, 2005&lt;sup&gt;76&lt;/sup&gt;; Allaart et al., 2006&lt;sup&gt;77&lt;/sup&gt;; Goekoop-Ruiterman, 2007&lt;sup&gt;83&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;85&lt;/sup&gt;; van der Kooij, 2009&lt;sup&gt;86&lt;/sup&gt;; Dirven et al., 2012&lt;sup&gt;87&lt;/sup&gt;; Dirven et al., 2013&lt;sup&gt;88&lt;/sup&gt;; Klarenbeek et al., 2010&lt;sup&gt;89&lt;/sup&gt;; Markusse et al., 2016&lt;sup&gt;90&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;91&lt;/sup&gt;; Klarenbeek et al., 2011&lt;sup&gt;92&lt;/sup&gt;; Markusse et al., 2014&lt;sup&gt;93&lt;/sup&gt;</td>
<td>BeST Study (continued)</td>
<td>Decrease in HAQ score, median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G5a: -0.38 (IQR, -0.63, 0.06)</td>
<td>G6a: -0.75 (IQR, -1.13, -0.25)</td>
<td>G5a vs. G6a: p&lt;0.001</td>
<td>G5b: -0.38 (IQR, 0.75, 0)</td>
<td>G6b: -0.63 (IQR, -1.13, -0.25)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF-36 PCS, improvement from baseline, mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 5.8</td>
<td>G2: 3.9</td>
<td>G3: 11.2</td>
<td>G4: 9.6</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF-36 MCS improvement from baseline, mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 2.1</td>
<td>G2: 2.5</td>
<td>G3: 0.4</td>
<td>G4: 3.1</td>
<td>p=0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Systolic Blood pressure (mm Hg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diastolic Blood pressure (mm Hg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Figure only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study</strong></td>
<td><strong>Interventions and Patient Characteristics</strong></td>
<td><strong>Baseline Disease and Treatment Characteristics</strong></td>
<td><strong>Health Outcomes</strong></td>
<td><strong>Adverse Events (%)</strong></td>
<td><strong>ROB Rating</strong></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td><strong>Characteristics</strong></td>
<td><strong>Study Population Summary</strong></td>
<td><strong>Interventions, dose:</strong></td>
<td><strong>Mean disease duration, mos:</strong></td>
<td><strong>No significant differences in efficacy between combination (MTX, SSZ) and single therapy (MTX or SSZ), only a trend favoring combination therapy, MTX and SSZ were comparable</strong></td>
<td><strong>Overall Events (%)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Author, yr:</strong> Haagsma, 1997</td>
<td>Patients meeting ACR criteria for RA with symptom duration &lt;1 yr, who were DMARD-naïve</td>
<td>G1: SSZ (1 g/day; max 3 g/day)</td>
<td>2.6-3.1</td>
<td>Overall:</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td><strong>Country, Setting:</strong> Netherlands, 1 academic and 6 peripheral clinics</td>
<td>G2: MTX (7.5 mg/wk; max 15 mg/wk)</td>
<td>Prior csDMARD use, %: 0</td>
<td></td>
<td>G1: 88.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong> RCT</td>
<td>G3: MTX (7.5 mg/wk; max 15 mg/wk) + SSZ (1 g/day; max 3 g/day)</td>
<td>Prior CS use, %: 0</td>
<td></td>
<td>G2: 77.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong> 105</td>
<td><strong>N:</strong> G1: 34</td>
<td><strong>MTX naive, %:</strong> 100</td>
<td></td>
<td>G3: 88.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2: 35</td>
<td><strong>Baseline DAS, mean:</strong> 4.6-5.0</td>
<td></td>
<td>SAEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G3: 36</td>
<td></td>
<td></td>
<td>G1: 8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Mean age, yrs:</strong> 54.9-57.0</td>
<td><strong>At 1 yr</strong></td>
<td></td>
<td>G2: 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAS mean change:</td>
<td></td>
<td>G3: 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: -1.6 (95% CI, -2.0 to -1.2)</td>
<td></td>
<td>Overall discontinuation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: -1.7 (95% CI, -2.0 to -1.4)</td>
<td></td>
<td>G1: 35.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G3: -1.9 (95% CI, -2.2 to -2.3)</td>
<td></td>
<td>G2: 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time to discontinuation in G1 &gt; G2, G3 (p=0.006)</td>
<td></td>
<td>G3: 16.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td>Haagsma, 1997&lt;sup&gt;23&lt;/sup&gt; (continued)</td>
<td>Sex, % female: 61.8-66.7</td>
<td>Baseline HAQ: 0.92-1.20</td>
<td>HAQ change from baseline: G1: -0.32 (95% CI, -0.53 to -0.10) G2: -0.46 (95% CI, -0.68 to -0.25) G3: -0.51 (95% CI, -0.76 to -0.26)</td>
<td>Discontinuation because of AEs G1: 26.5 G2: 5.7 G3: 13.9</td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>1 yr</td>
<td>Race, % white: NR</td>
<td>Seropositive (RF or CCP) (%): RF+: 94.2-97.1</td>
<td>N of pts with a response according to ACR criteria at end of study: G1: 25 G2: 25 G3: 28</td>
<td>Discontinuation because of lack of efficacy G1: 8.8 G2: 0.0 G3: 2.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline Sharp score, mean: NR</td>
<td>Patient adherence &gt;90% for all patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosive disease, %: NR</td>
<td>AEs possibly/probably related to treatment G1: 47.1 G2: 31.4 G3: 63.9 G3 &gt; G1, G2 (p=0.023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cardiovascular Events (Dyspnea): G1: 5.9 G2: 0 G3: 5.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nausea: G1: 29.4 G2: 25.7 G3: 63.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>URTI G1: 17.6 G2: 20.0 G3: 27.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Heimans et al., 2013; Heimans et al., 2014; Heimans et al., 2016 IMPROVED</td>
<td>Adults aged ≥ 18 yrs RA (fulfilling ACR and EULAR criteria for RA with symptom duration ≤ 2 yrs) or UA (≥ 1 joint with clinical synovitis and ≥ 1 other painful joint, clinically suspected as due to early RA regardless of symptom duration), DAS ≥ 1.6, no prior antirheumatic therapy, and for whom MTX 25 mg/wk with PRED: 60 mg/day tapered to 7.5 mg/day had not lowered their DAS28 to ≤ 1.6 during the first 4 mos of disease treatment</td>
<td>Interventions, dose: G1: • MTX: 25 mg/wk • PRED: 7.5 mg/day • HCQ: 400 mg/day • SSZ: 2000 mg/day • PRED, HCQ, SSZ stopped if remission achieved at 8 mos; switched to 25 mg/wk MTX and 40 mg/every other wk ADA if remission not achieved at 8 mos (12 mos of treatment) • MTX stopped if remission remained at 12 mos (16 mos of treatment) G2: • MTX: 25 mg/wk • ADA: 40mg every other wk ADA tapered if remission achieved at 8 mos; ADA increased to 40 mg/wk if remission not achieved at 8 mos (12 mos of treatment) non-MTX drugs were stopped/tapered in patients who achieved remission after 8 mos; MTX was stopped if remission remained 4 mos later. Patients in G1 that did not achieve remission at 8 mos received ADA therapy instead. Patients in G2 that did not achieve remission at 8 mos received an increased dose of 40 mg/wk of ADA.</td>
<td>Median disease duration, wks: G1: 22 (IQR, 9-40) G2: 21 (IQR, 8-29) Overall: NR</td>
<td>At 2 yrs (20 mos after randomization)</td>
<td>Overall AEs in yr 2: G1: 63.9 G2: 66.7</td>
<td>High</td>
</tr>
</tbody>
</table>
| **Country, Clinical Setting:** Netherlands, multicenter (12 hospitals) | 4-mos DAS, mean: G1: 2.49 (SD, 0.63) G2: 2.57 (SD, 0.68) Overall: NR | DAS disease activity, mean: G1: 2.02 (SD, 0.70) G2: 1.92 (SD, 0.85) p=0.45 | Overall
DAS remission (< 1.6), %: G1: 26.5 G2: 30.8 p=0.76 | **4-mos HAQ, mean:** G1: 0.86 (SD, 0.57) G2: 0.88 (SD, 0.57) Overall: NR | **ACR response, %:** NR | **SAEs in yr 2:** G1: 6 G2: 10.2 | **Overall discontinuation:** NR |
<p>| <strong>Study Design:</strong> RCT | MTX naïve, %: 0.0 | mTSS score, progression (increase ≥ 0.5), %: G1: 10.8 G2: 6.4 p=0.31 | <strong>HAQ, mean:</strong> G1: 0.90 (SD, 0.66) G2: 0.83 (SD, 0.67) | SF-36: NR | Discontinuation because of lack of efficacy: NR |
| | MTX inadequate responders, %: 100 | <strong>12 mos (8 mos after randomization)</strong> | <strong>Biologic non-responders, %:</strong> NR | <strong>ACPA positive, %:</strong> G1: 48.2 G2: 46.2 Overall: NR | Overall AEs, %: G1: 74 G2: 68 P = 0.41 | <strong>Discontinuation because of lack of AEs:</strong> NR |
| | Prior DMARD use, %: 100 | <strong>DAS disease activity, mean:</strong> G1: 2.07 (SD, 0.89) G2: 1.77 (SD, 0.90) p=0.04 | <strong>Specific AEs:</strong> Increased liver enzymes: G1: 8.4 G2: 3.8 | <strong>Between 4 mos (randomization) and yr 1</strong> | <strong>Patient adherence:</strong> NR | <strong>SAEs, %:</strong> G1: 8.4 G2: 9 | <strong>Discontinuation because of AEs:</strong> NR |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Heimans et al., 2013; Heimans et al., 2014; Heimans et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall discontinuation, %: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation because of AEs, %: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation because of lack of efficacy, %: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Patient adherence: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Specific AEs:</strong> Increased liver enzymes: G1: 6 G2: 11.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>URTI: G1: 4.8 G2: 10.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nausea: G1: 7.2 G2: 6.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Headache: G1: 8.4 G2: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dizziness: G1: 1.2 G2: 0</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr,</strong> Study Name: Heimans et al., 2013; Heimans et al., 2016</td>
<td><a href="#"><strong>Overall N:</strong> 161</a></td>
<td>N: G1: 83 G2: 78</td>
<td>Baseline mTSS score, median: G1: 0 (IQR, 0-0.5) G2: 0 (IQR, 0-0)</td>
<td>DAS remission (&lt; 1.6), % G1: 25.3 G2: 41.0 p=0.01</td>
<td>Pneumonia or bronchitis G1: 3.6 G2: 1.3</td>
<td></td>
</tr>
<tr>
<td>Study Duration: 2 yrs</td>
<td></td>
<td>Mean age, yrs: 49-51</td>
<td>Erosive disease, %: G1: 12.0 G2: 16.7 Overall: NR</td>
<td>Total SHS, median (IQR): G1: 0 (0.0-0.5) G2: 0 (0-0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female: 74-77</td>
<td></td>
<td>SHS progression median (IQR): G1: 0 (0-0) G2: 0 (0-0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td></td>
<td>HAQ, mean: G1: 0.87 (SD, 0.66) G2: 0.81 (SD, 0.66) p=0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: NR</td>
<td></td>
<td>SF-36: Mental component, mean: G1: 50.5 (SD, 10.3) G2: 50.5 (SD, 10.1) p=0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td></td>
<td>Physical component, mean: G1: 39.9 (SD, 10.3) G2: 43.0 (SD, 11.4) p=0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pain (visual analog scale), mean: G1: 38 (SD, 28) G2: 28 (SD, 25) p=0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VAS global health (mm), mean: G1: 33 (SD 23) G2: 27 (SD 20)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics

**Author, yr, Study Name:**
Heimans et al., 2013;
Heimans et al., 2016120

**IMPROVED (continued)**

### Study Population Summary

### Interventions and Patient Characteristics

### Baseline Disease and Treatment Characteristics

### Health Outcomes

<table>
<thead>
<tr>
<th>Erosive, %</th>
<th>8 mos (4 mos after randomization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 15</td>
<td>DAS disease activity, mean:</td>
</tr>
<tr>
<td>G2: 15</td>
<td>G1: 1.97 (SD, 0.87)</td>
</tr>
<tr>
<td></td>
<td>G2: 2.01 (SD, 0.91)</td>
</tr>
<tr>
<td></td>
<td>p=0.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HAQ, mean:</th>
<th>SF-36:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 0.74 (SD, 0.61)</td>
<td>Mental component, mean:</td>
</tr>
<tr>
<td>G2: 0.81 (SD, 0.64)</td>
<td>G1: 46.6 (SD, 17.9)</td>
</tr>
<tr>
<td>p=0.51</td>
<td>G2: 48.7 (SD, 10.3)</td>
</tr>
<tr>
<td></td>
<td>p=0.85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical component, mean:</th>
<th>Pain (visual analog scale), mean:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 42.8 (SD, 10.9)</td>
<td>G1: 35 (SD, 26)</td>
</tr>
<tr>
<td>G2: 42.5 (SD, 11.0)</td>
<td>G2: 31 (SD, 25)</td>
</tr>
<tr>
<td>p=0.10</td>
<td>p=0.36</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Hellgren et al., 2017; SRQ Register analysis</td>
<td>Patients meeting 1987 revised ACR criteria for RA between 1997 and 2012 with disease duration &lt;1 yr between first RA symptom and diagnosis</td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> Sweden, outpatient</td>
<td><strong>Overall N:</strong> 12,656</td>
</tr>
<tr>
<td><strong>Study Design:</strong> Single-arm study</td>
<td><strong>G1:</strong> 55 • G1a: 40 • G1b: 22 • G1c: 12</td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 15 years</td>
<td>MTX naive: 100</td>
</tr>
<tr>
<td></td>
<td>Biologic non-responders: NA</td>
</tr>
<tr>
<td></td>
<td>Erosive disease, %: NR</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Hellgren et al., 2017; SRQ Register analysis (continued)</td>
<td>G2: 11,638</td>
</tr>
<tr>
<td>Mean age, yrs:</td>
<td>58</td>
</tr>
<tr>
<td>Sex, % female:</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
| **Author, yr, Study Name:** | Horslev-Petersen et al., 2014;164 | Adults (aged ≥ 18 yrs) fulfilling ACR criteria for RA with disease duration < 6 mos, moderate to severe RA defined as DAS28 CRP > 3.2, no prior DMARD use, and no treatment with glucocorticoids within last 4 wks | **Interventions, dose:** G1: - MTX: 7.5 mg/wk, increased to 15 mg/wk after 1 mo and 20 mg/wk after 2 mos - ADA: 40 mg every other wk (subcutaneous) G2: - MTX: 7.5 mg/wk, increased to 15 mg/wk after 1 mo and 20 mg/wk after 2 mos - Placebo ≤ 4 swollen joints observed at each visit (total 7) were injected with triamcinolone hexacetonide (40 mg/ml, 0.5-2 ml/joint); if unacceptable disease activity persisted at 3 mos or thereafter (defined as either DAS28 CRP ≥ 3.2 and ≥ 1 swollen joint or intra-articular injection of 4 ml triamcinolone was given monthly for 3 consecutive mos), 200 mg/day HCQ and 2,000 mg/day SSZ were added; if LDA was not achieved within an additional 3 mos, | **Mean disease duration, days:** G1: 83-88 **Baseline DAS28 CRP, mean:** 5.5-5.6 **Baseline HAQ, median:** 1.0-1.1 **MTX naive:** 100 **Prior csDMARD use, %:** 0 **MTX inadequate responders:** NA **Biologic non-responders:** NA **Prior CS use, %:** 0 **RF seropositive, %:** 72.0 | At 2 yrs (1 yr after stopping ADA) DAS28 CRP disease activity, median change: G1: -3.1 (5/95% range: -1.0 to 5.7) G2: -3.1 (5/95% range: -1.3 to 5.1) p=0.72 | Overall AEs: NR
SAEs: At 1 yr (%): G1: 15.7 G2: 11.0 During yr 2 (n): G1: 4 G2: 11 Overall discontinuation: At 1 yr G1: 9.0 G2: 12.1 At 2 yrs G1: 10.1 G2: 16.5 Discontinuation because of AEs: At 1 yr G1: 2.2 G2: 1.1 At 2 yrs No discontinuations due to AEs | Medium |
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name:</td>
<td>Horslev-Petersen et al., 2014;</td>
<td>ADA/placebo was discontinued, and the patient was considered a non-responder and prescribed open-label non-ADA biologics</td>
<td>anti-CCP seropositive, %: 65.1</td>
<td>Sharp score Median change:</td>
<td>Discontinuation because of lack of efficacy:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Axelsen et al., 2015;</td>
<td>N: G1: 89 G2: 91</td>
<td>Baseline Sharp score, median: 4.3-4.5</td>
<td>G1: 1.05 G2: 2.63 p=0.12</td>
<td>At 1 yr G1: 3.4 G2: 2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ørnbjerg et al., 2017;</td>
<td>Mean age, yrs: G1: 56.2 (25.8-77.6) G2: 54.2 (29.3-76.7) Overall: NR</td>
<td>Erosive disease, %: 53</td>
<td>No radiographic progression (change ≤ 0), %:</td>
<td>At 2 yrs No discontinuations due to lack of efficacy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Horslev-Petersen et al.,</td>
<td>Sex, % female: 66.0</td>
<td>HAQ Median change:</td>
<td>G1: 64 G2: 51 p=0.81</td>
<td>Discontinuation due to patient request/non-compliance:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ammitzboll et al., 2013;</td>
<td>Race, % white: NR</td>
<td>G1: -0.9 (5/95% range: 0.3 to -2.5)</td>
<td>Overall: 4.4</td>
<td>At 1 yr Overall: 4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OPERA (continued)</td>
<td>Race, % black: NR</td>
<td>G2: -0.6 (5/95% range: 0.5 to -1.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td>p=0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Response (&lt; 0.5), %:</td>
<td>G1: 70 G2: 64 p=0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36 Mental component, median:</td>
<td>G1: 56 (5/95% range: 36 to 62) G2: 56 (5/95% range: 34 to 64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 56 (5/95% range: 23 to 57)</td>
<td>p=0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Physical component, median:</td>
<td>G1: 46 (5/95% range: 23 to 57) G2: 45 (5/95% range: 22 to 56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: -36 (5/95% range: 13 to -88) G2: -31 (5/95% range: 6 to -80)</td>
<td>p=0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pain (VAS), median change:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p=0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fatigue (VAS), median change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td>G1: -32 (5/95% range: 2 to -79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -22 (5/95% range: 34 to -75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>At 1 yr</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 CRP disease activity, median:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2.0 (5/95% range: 1.7 to 5.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 2.6 (5/95% range: 1.7 to 4.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR20 response, %:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR50 response, %:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR70 response, %:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>DAS28 CRP remission (&lt; 2.6),%:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.0008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Sharp score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Median change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 1.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td></td>
<td></td>
<td></td>
<td>No radiographic progression (change ≤ 0), %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horslev-Petersen et al., 2014;</td>
<td></td>
<td></td>
<td></td>
<td>G1: 67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016;</td>
<td></td>
<td></td>
<td></td>
<td>G2: 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axelsen et al., 2015;</td>
<td></td>
<td></td>
<td></td>
<td>p=0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ørnbjerg et al., 2017;</td>
<td></td>
<td></td>
<td></td>
<td>HAQ, median change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horslev-Petersen et al., 2016;</td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.88 (5/95% range: -2.46 to 0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammitzboll et al., 2013;</td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.63 (5/95% range: -1.82 to 0.38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERA (continued)</td>
<td></td>
<td></td>
<td></td>
<td>p=0.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td></td>
<td></td>
<td></td>
<td>Mental component, median change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 5.5 (5/95% range: -8.5 to 20.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 4.3 (5/95% range: -9.3 to 27.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical component, median change:</td>
<td></td>
<td></td>
<td></td>
<td>G1: 13.2 (5/95% range: -2.3 to 33.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 10.6 (5/95% range: -11.2 to 22.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain (visual analogue scale), median:</td>
<td></td>
<td></td>
<td></td>
<td>G1: 7 (5/95% range: 0 to 64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 20 (5/95% range: 0 to 71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue (visual analogue scale), median:</td>
<td></td>
<td></td>
<td></td>
<td>G1: 16 (5/95% range: 0 to 81)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 20 (5/95% range: 0 to 84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D, median change from baseline (5th/95th percentile ranges)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horslev-Petersen et al., 2014;</td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.22 (-0.05 to 0.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axelsen et al., 2015;</td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.20 (-0.06 to 0.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ørnbjerg et al., 2017;</td>
<td></td>
<td></td>
<td></td>
<td>(p=0.095)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horslev-Petersen et al., 2016;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammitzboll et al., 2013;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERA (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics
- **Author, yr, Study Name:** Kavanaugh et al., 2013; Smolen et al., 2014; Emery et al., 2016
- **Study Name:** OPTIMA
- **Country, Clinical Setting:** Multiple countries, 161 sites (Academic hospitals, research centers, private practices and rheumatology clinics)
- **Study Design:** RCT
- **Overall N:** 1032
- **Study Duration:** 78 wks (open label after 26 wks)

### Study Population Summary
- Patients aged ≥18 years with RA diagnosis based on ACR criteria, with disease duration <1 yr. All patients were MTX and biological DMARD naïve.

### Interventions and Patient Characteristics
- **Interventions, dose:**
  - G1: ADA 40 mg/every other wk + MTX 7.5 mg/wk (maximum of 20 mg/wk by wk 8)
  - G1a: Randomized to placebo + MTX (ADA withdrawal)
  - G1b: Randomized to ADA + MTX (ADA continuation)
  - G1c: Open-label ADA + MTX (ADA carry-over)
  - G2: Placebo + MTX
  - G2a: Continued masked placebo + MTX monotherapy
  - G2b: Open-label ADA + MTX (ADA rescue)
- **MTX naïve, %:** 100
- **MTX inadequate responders:** 0
- **RF Seropositive (%):** 88.3
- **Sharp score (modified total), mean:** 11.2-11.8
- **Patients with ≥1 erosion, %:** 83.4

### Baseline Disease and Treatment Characteristics
- **Mean age, yrs:** 50.4-50.7
- **Sex, % female:** 74
- **Race, % white:** 89.5
- **Mean disease duration, mos:** 4.0-4.5 mos
- **Baseline DAS, mean:** 6.0
- **Baseline HAQ-DI (0-3), mean:** 1.60-1.61
- **MTX naïve, %:** 100
- **MTX inadequate responders:** 0
- **Biologic non-responders:** NR
- **ACR20 response, %**
  - G1a: 94.1
  - G1b: 95.2
  - G1c: NR
  - G2a: 91.1
  - G2b: 83
- **ACR50 response, %**
  - G1a: 80.4
  - G1b: 88.6
  - G1c: NR
  - G2a: 76.8
  - G2b: 63
- **ACR70 response, %**
  - G1a: 64.7
  - G1b: 77.1
  - G1c: NR
  - G2a: 61.6
  - G2b: 43
- **DAS28 (CRP) <2.6 remission, %**
  - G1a: 66.3
  - G1b: 85.7
  - G1c: NR
  - G2a: 67.9
  - G2b: 44
- **At 78 wks DAS28 <3.2 % achieving LDA**
  - G1a: 81.2
  - G1b: 91.4
  - G1c: NR
  - G2a: 81.3
  - G2b: 60

### Health Outcomes
- **Overall:**
  - Period 1:
    - G1: 73.6
    - G2: 71.2
  - Period 2:
    - G1a: 77.5
    - G1b: 71.4
    - G1c: 76.8
    - G2a: 74.1
    - G2b: 77.6
- **SAEs**
  - Period 1:
    - G1: 7.2
    - G2: 6.2
  - Period 2:
    - G1a: 10.8
    - G1b: 11.4
    - G1c: 6.9
    - G2a: 8.0
    - G2b: 9.2
- **Overall discontinuation**
  - G1 total: 22.3
  - G2 total: 24.2
  - Period 1:
    - G1: 10
    - G2: 11
  - Period 2:
    - G1a: 12.7
    - G1b: 9.5
    - G1c: 16.6
    - G2a: 13.4
    - G2b: 15.2

### ROB Rating
- **Low**
### Study Characteristics

**Study Population Summary**

**Interventions and Patient Characteristics**

**Baseline Disease and Treatment Characteristics**

**Health Outcomes**

<table>
<thead>
<tr>
<th>Author, yr, Study Name: Kavanaugh et al., 2013;17 Smolen et al., 2014;151 Emery et al., 2016152 OPTIMA (continued)</th>
<th>Change from baseline in mTSS ≤0.5 (%)</th>
<th>Discontinuation because of AEs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G1a: 80.6</td>
<td>Overall:</td>
</tr>
<tr>
<td></td>
<td>G1b: 89.3</td>
<td>G1: 8.9</td>
</tr>
<tr>
<td></td>
<td>G1c: NR</td>
<td>G2: 7.9</td>
</tr>
<tr>
<td></td>
<td>G2a: 78.0</td>
<td>Period 2</td>
</tr>
<tr>
<td></td>
<td>G2b: Figure only</td>
<td>G1a: 6.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1b: 2.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1c: 6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2a: 5.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2b: 5.7</td>
</tr>
</tbody>
</table>

**Adverse Events (%)**

<table>
<thead>
<tr>
<th>ROB Rating</th>
<th>Patient adherence</th>
<th>NR</th>
</tr>
</thead>
</table>

**Bronchitis**

| | G1a: 0 | G1b: 0 |
| | G1c: 0 | G2a: 0.9 |
| | G2b: 0 |

**Dizziness**

| | G1a: 1.0 | G1b: 0 |
| | G1c: 0 | G2a: 0 |
| | G2b: 0 |

**Discontinuation**

<table>
<thead>
<tr>
<th></th>
<th>HAQ-DI (0-3), mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G1a: 0.38 (CI 0.27 to 0.50)</td>
</tr>
<tr>
<td></td>
<td>G1b: 0.35 (CI 0.25 to 0.45)</td>
</tr>
<tr>
<td></td>
<td>G1c: 0.89 (CI 0.81 to 0.98)</td>
</tr>
<tr>
<td></td>
<td>G2a: 0.39 (CI 0.29 to 0.48)</td>
</tr>
<tr>
<td></td>
<td>G2b: 0.76 (CI 0.69 to 0.83)</td>
</tr>
</tbody>
</table>

**SF-36**

<table>
<thead>
<tr>
<th></th>
<th>NR</th>
</tr>
</thead>
</table>

**At 26 wks**

<table>
<thead>
<tr>
<th>DAS, % achieving LDA target</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 47</td>
</tr>
<tr>
<td>G2: 26 (p&lt;0.001)</td>
</tr>
</tbody>
</table>

**ACR20 response, %**

| | G1: 70 |
| --- | G2: 57 (p<0.001) |

**ACR50 response, %**

| | G1: 52 |
| --- | G2: 34 (p<0.001) |

**ACR70 response, %**

<p>| | G1: 35 |
| --- | G2: 17 (p&lt;0.001) |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Kavanaugh et al., 2013; Smolen et al., 2014; Emery et al., 2016</td>
<td></td>
<td></td>
<td></td>
<td>DAS &lt;2.6 remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 34</td>
<td>G2: 17 (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mTSS mean change</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.15</td>
<td>G2: 0.96 (P &lt;0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI, mean value</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.7</td>
<td>G2: 0.9 (P &lt;0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal function (HAQ-DI &lt;0.5), %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 40</td>
<td>G2: 28 (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPAI activity impairment, mean % change from baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 32.0</td>
<td>G2: 23.7 (p=0.0071)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPAI presenteeism (performance while at work owing to RA), mean % change from baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 24.6</td>
<td>G2: 17.1 (p=0.0253)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WPAI overall work impairment, mean % change from baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 27.3</td>
<td>G2: 18.3 (p=0.0105)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Kavanaugh et al., 2013;</td>
<td></td>
<td></td>
<td>At 22 wks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Smolen et al., 2014;</td>
<td></td>
<td></td>
<td>DAS, % achieving LDA target</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emery et al., 2016</td>
<td></td>
<td></td>
<td>G1: 44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OPTIMA (continued)</td>
<td></td>
<td></td>
<td>G2: 24 (P &lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS, mean change in modified total score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.96 (P &lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI, mean value</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.9 (P &lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Kellner et al., 2010</td>
<td>Adult patients with early RA (defined by a max disease duration of 1 year since diagnosis) were eligible if the investigator was convinced that they might profit from treatment with LEF and if they did not show any contraindications. The physician's decision for LEF treatment was based on patient's condition and independent of study documentation.</td>
<td><strong>Interventions, dose:</strong> Recommended loading dose was LEF, 100 mg/d. Maintenance dose was LEF 20 mg/d in 91.6% of patients and 10 mg/d in 8.4% of patients. 61.7% were concomitantly treated with corticosteroids, and in 27.5% of patients additional DMARDs (most often MTX, 22.2%) were used.</td>
<td><strong>Median time since RA diagnosis, mos:</strong> 4.0</td>
<td><strong>N/A</strong></td>
<td><strong>Overall:</strong> 10.8</td>
<td></td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> Germany, 174 centers</td>
<td></td>
<td></td>
<td><strong>Mean disease duration, mos (SD):</strong> 7.5 (SD, 15.8)</td>
<td></td>
<td><strong>SAEs (“Serious adverse drug reactions”):</strong> 1.2</td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong> Observational (only single arm eligible)</td>
<td></td>
<td></td>
<td><strong>Baseline DAS, mean:</strong> 5.7 (SD, 1.2)</td>
<td></td>
<td><strong>Overall discontinuation:</strong> 11.1</td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong> 334</td>
<td></td>
<td></td>
<td><strong>Baseline HAQ-DI, mean:</strong> 1.37 (SD, 0.7)</td>
<td></td>
<td><strong>Discontinuation because of AEs:</strong> 6.3</td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 25.5 wks observance on average</td>
<td></td>
<td></td>
<td><strong>MTX naive:</strong> 58.1</td>
<td></td>
<td><strong>Patient adherence:</strong> NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>MTX inadequate responders:</strong> NR</td>
<td></td>
<td><strong>Specific AEs:</strong> NA (specific AEs for head-to-head trials only)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Kellner et al., 2010[19] (continued)</td>
<td></td>
<td></td>
<td>CCP seropositive (%)</td>
<td>60.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline Sharp score, mean:</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosive disease, %:</td>
<td>45.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Leirisalo-Repo et al., 2013</td>
<td>Patients aged 18-60, fulfilling ACR criteria for RA, DMARD naive, and not permanently work disabled or retired. All had active disease (≥6 swollen joints/≥6 tender joints) and either early morning stiffness ≥45 min, ESR rate ≥30 mm/h or CRP ≥20 mg/l</td>
<td><strong>Interventions, dose:</strong> G1: “FIN-RACo” (MTX + SSZ + HCQ + PRED) + IFX (3 mg/kg from wks 4-26) G2: “FIN-RACo” (MTX + SSZ + HCQ + PRED) + Placebo (from wks 4-26) FIN-RACo: Regimen consisting of:  - MTX: Starting at 10 mg/wk, 15 mg/wk at wk 4, 20 mg/wk at wk 10, 25 mg/wk from wk 14  - SSZ: Starting at 1 g/d, 2 g/d at 2 wks, 1-2 g/d from wk 4  - HCQ: 35 mg/kg/wk from start through study duration  - PRED: 7.5 mg/d from start through study duration  - Acid folic with MTX (5 mg/wk), Calcium (1000 mg/d), and Vitamin D3 (800 IU/d) throughout study IFX: Received at wks 4, 6, 10, 18, and 26</td>
<td><strong>Median disease duration, mos (IQR):</strong> 4 (IQR, 2, 6) <strong>Baseline DAS, mean:</strong> 5.5-5.6 <strong>Baseline HAQ, mean:</strong> 0.9-1.1 <strong>MTX naive:</strong> 100 <strong>MTX inadequate responders:</strong> 0 <strong>Biologic non-responders:</strong> NR <strong>RF seropositive (%):</strong> 76 <strong>Baseline Sharp score, mean:</strong> 2.0-2.8 <strong>Erosive disease, %:</strong> 37</td>
<td><strong>At 5 years followup DAS disease activity, mean (SD):</strong> G1: 2.0 (SD, 1.2) G2: 1.7 (SD, 0.9) p=0.33 <strong>ACR20 response, %</strong> NR <strong>ACR50 response, %</strong> NR <strong>ACR70 response, %</strong> NR <strong>ACR strict remission, % (95% CI):</strong> G1: 60 (CI 44 to 74) G2: 61 (CI 45 to 75) p=0.93 <strong>DAS remission, % (95% CI):</strong> G1: 84 (CI 71 to 94) G2: 89 (CI 76 to 96) p=0.51 <strong>SHS scores, mean (SD):</strong> G1: 4.3 (SD, 7.6) G2: 5.3 (SD, 7.3) p=0.54 <strong>HAQ, median (IQR):</strong> G1: 0 (IQR, 0.0-0.1) G2: 0 (IQR, 0.0-0.0) p=0.39</td>
<td><strong>Overall:</strong> Year 5: G1: 91.3 G2: 97.9 Year 2: G1: 90 G2: 96 SAEs Year 5: G1: 8.7 G2: 10.6 p=0.99 Year 2: G1: 6 G2: 8 Overall discontinuation Year 5: G1: 10 G2: 6.1 Year 2: G1: 8 G2: 8.2 Discontinuation because of AEs Year 2: G1: 2 G2: 0</td>
<td>Low</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
</tbody>
</table>
| Author, yr, Study Name:  
Leirisalo-Repo et al, 2013  
Rantalaiho et al., 2014  
Kuusalo et al., 2015  
NEO-RACo (continued) | Sex, % female:  
67 | | At 2 years  
DAS disease activity  
NR  
ACR20 response, %  
NR  
ACR50 response, % (95% CI)  
G1: 96 (CI 86 to 100)  
G2: 92 (CI 80 to 98)  
p=0.436  
ACR70 response, % (95% CI)  
G1: 86 (CI 73 to 94)  
G2: 71 (CI 57 to 83)  
p=0.090  
ACR modified remission, % (95% CI)  
G1: 66 (CI 51 to 81)  
G2: 53 (CI 38 to 67)  
p=0.19  
DAS28 remission, % (95% CI)  
G1: 82 (CI 72 to 93)  
G2: 82 (CI 71 to 93)  
SHS score, mean change from baseline (95% CI)  
G1: -0.2 (CI -1.2 to 0.4)  
G2: 1.4 (CI 0.8 to 2.3)  
p=0.0058 | Patient adherence  
Year 2:  
95% of patients sufficiently complied with the study protocol  
Year 5:  
NR  
Specific AEs  
GI symptoms: 56% vs. 61%  
Respiratory: 56% vs. 67%  
Elevated liver enzymes: 12% vs. 16% |
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Leirisalo-Repo et al, 2013</td>
<td>Patients with an RA diagnosis, disease duration ≤ 1 yr, and no prior DMARD use who were enrolled in the NOR-DMARD register and starting treatment with SSZ or MTX as monotherapies</td>
<td><strong>Interventions, dose:</strong></td>
<td><strong>Baseline DAS28, mean:</strong></td>
<td><strong>At 6 months</strong></td>
<td><strong>Overall AEs:</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: SSZ (median dose of 2.0 g [IQR 2.0-2.0] at all timepoints)</td>
<td>G2: MTX (median dose of 10 mg [IQR 7.5-15.0] at baseline, 15 mg [IQR 12.5-15.0] at 3 mos, 15 mg [IQR 12.5-20.0] at 6 mos)</td>
<td>G1: 4.38 (SD, 1.35) G2: 5.00 (SD, 1.34)</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Rantalaiho et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR20 response, % - LUNDEX</strong></td>
<td>G1: 20.8 G2: 44.5 (p=NA) <strong>ACR50 response, % - LUNDEX</strong></td>
</tr>
<tr>
<td></td>
<td>Kuusalo et al., 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NEO-RACo (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Lie et al., 2012</td>
<td>Patients with an RA diagnosis, disease duration ≤ 1 yr, and no prior DMARD use who were enrolled in the NOR-DMARD register and starting treatment with SSZ or MTX as monotherapies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NOR-DMARD analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong></td>
<td>Norway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td>register-based longitudinal observational study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong></td>
<td>1,102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong></td>
<td>3 yrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Lie et al., 2012\textsuperscript{28}</td>
<td>NOR-DMARD analysis (continued)</td>
<td>Biologic non-responders, %: NR</td>
<td>Mean modified HAQ (MHAQ) change from baseline (SD)</td>
<td>Nausea: G1: 13.1 G2: 18.9 p&lt;0.07</td>
<td>G1: 13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RF seropositive, %: G1: 50.3 G2: 61.4 Overall: NR</td>
<td>G1: -0.13 (0.45) G2: -0.26 (0.48) (p=0.002 from t-test; p=0.05 from ANCOVA adjusted for propensity score quintile; p=0.13 from ANCOVA adjusted for propensity score quintile and physician global VAS)</td>
<td>Abdominal pain: G1: 8.0 G2: 4.1 p&lt;0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline Sharp score, mean: NR Erosive disease, %: NR</td>
<td>Mean SF-36 PCS change from baseline (SD)</td>
<td>Rash: G1: 9.1 G2: 2.7 p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 4.0 (8.5) G2: 5.4 (9.8) (p=0.11 from t-test; p=0.26 from ANCOVA adjusted for propensity score quintile; p=0.42 from ANCOVA adjusted for propensity score quintile and physician global VAS)</td>
<td>Hair loss: G1: 1.1 G2: 5.1 p&lt;0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stomatitis: G1: 0.6 G2: 4.4 p&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Lie et al., 2012</td>
<td>NOR-DMARD (continued)</td>
<td></td>
<td><strong>Mean Pain VAS change from baseline (SD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -9.2 (23.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -14.7 (26.9) (p=0.02 from t-test; p=0.24 from ANCOVA adjusted for propensity score quintile; p=0.41 from ANCOVA adjusted for propensity score quintile and physician global VAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Mean Fatigue VAS change from baseline (SD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.4 (28.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -4.4 (29.6) (p=0.13 from t-test; p=0.21 from ANCOVA adjusted for propensity score quintile; p=0.24 from ANCOVA adjusted for propensity score quintile and physician global VAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Data not abstracted for patient matching analysis (according to RF status and baseline DAS28) because only unadjusted comparisons of their data were performed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>At 3 months ACR20 response, % - LUNDEX</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 18.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 47.4 (p=NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>ACR50 response, % - LUNDEX</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 5.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 21.3 (p=NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Lie et al., 2012 NOR-DMARD analysis (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, % - LUNDEX</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 14.0 (p=NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission (&lt;2.6), % - LUNDEX</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 14.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 25.6 (p=NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Adults (aged ≥ 18 yrs) fulfilling ACR criteria for RA &lt; 6 mos, DAS28 &gt; 3.2, no prior DMARD or CS use</td>
<td>Interventions, dose:</td>
<td>Mean disease duration, mos:</td>
<td>At 24 wks</td>
<td>Overall AEs:</td>
<td>Medium</td>
</tr>
<tr>
<td>Marcora et al., 2006</td>
<td></td>
<td>G1:</td>
<td>NR</td>
<td>DAS28 disease activity, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong></td>
<td></td>
<td>ETN: 25 mg twice/wk (subcutaneous)</td>
<td></td>
<td>G1: 3.2 (SD 1.5)</td>
<td>G1: 0.0</td>
<td></td>
</tr>
<tr>
<td>United Kingdom, hospital outpatient rheumatology clinic</td>
<td></td>
<td>G2: MTX: 7.5 mg/wk for 1 mo, increased to max 15 mg/wk in mo 2 and 20 mg/wk in mo 4 if necessary (oral), with 10 mg/wk folic acid</td>
<td></td>
<td>G2: 3.1 (SD 1.4)</td>
<td>G2: 0.0</td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td>N:</td>
<td></td>
<td></td>
<td>Treatment x time: P = 0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td>G1: 12</td>
<td>MTX naive, %: 100</td>
<td></td>
<td>Time: P &lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong></td>
<td>G2: 14</td>
<td>MTX inadequate responders, %: NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>Biologic non-responders, %: NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong></td>
<td>Race, % white: 75.0 (of 24)</td>
<td>RF seropositive, %: 58.3 (of 24)</td>
<td></td>
<td>Sharp score, mean: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 mos</td>
<td>Race, % black: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean: G1: 1.0 (SD 0.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.6 (SD 0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Treatment x time: P = 0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time: P &lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 12 wks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 disease activity, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 3.8 (SD 1.5)</td>
<td>G1: 1.2 (SD 0.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 3.4 (SD 1.2)</td>
<td>G2: 0.6 (SD 0.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1:</td>
<td>G1: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2:</td>
<td>G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to AEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to lack of efficacy:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific AEs: Injection site reaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 8.3</td>
<td>G1: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.0</td>
<td>G2: 0.0</td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics

**Author, yr, Study Name:** McWilliams et al., 2013

**Study Population Summary:** Patients recruited after diagnosis of RA by rheumatologist. People whose diagnosis subsequently changed were removed from study database. Data were analyzed for patients who had been recruited prior to July 2009, ≥ 2 years before data retrieval for this analysis, and who had commenced DMARDs before visit 4.

**Country, Clinical Setting:** UK and Eire, 22 outpatient centers

**Study Design:** Observational, (Retrospective cohort)

**Overall N:** 766

**Study Duration:** 2 yrs

### Interventions and Patient Characteristics

**Interventions, dose:**
- **G1:** Initial DMARD regimen of SSZ monotherapy
- **G2:** Initial DMARD regimen of MTX monotherapy
- **G3:** Initial DMARD regimen of MTX + SSZ + HCQ triple therapy

**N:**
- G1: 273
- G2: 336
- G3: 52

**Mean age, yrs:** 56-58

**Sex, % female:** 65-72

**Race, % white:** NR

### Baseline Disease and Treatment Characteristics

**Baseline Disease duration, mos (IQR):**
- 6 mos (IQR, 4-12)

**Baseline DAS28, median (IQR):**
- 5.8 (IQR, 4.6-7.0)

**Baseline HAQ, median (IQR):**
- 1.1 (IQR, 0.6-1.8)

**MTX naive:**
- 100

**MTX inadequate responders:**
- NR

**Prior csDMARD use, %**
- 0

**Prior CS use, %**
- 16-17

**Biologic non-responders:**
- NR

**Seropositive (%):**
- 61-62

### Health Outcomes

**Adverse Events (%)**

<table>
<thead>
<tr>
<th>G1</th>
<th>G2</th>
<th>G3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changed DMARD, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>36</td>
<td>4</td>
</tr>
</tbody>
</table>

**Heat-to-head analysis, comparing G1 and G2:**
- MTX is favored as initial DMARD (aOR = 0.41 (0.28-0.60), p<0.001)

**Changed DMARD due to adverse drug reaction, %**
- G1: 59
- G2: 23
- G3: 2

**Note:** sensitivity analyses only including participants who satisfied ACR 1987 classification criteria for RA did not affect statistical associations between baseline factors and DMARD change (data not shown).

**Adverse Events (%)**

- **ROB Rating:** High
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: McWilliams et al., 2013</td>
<td></td>
<td>Baseline Sharp score, mean: NR</td>
<td>Erosions, %: 26-47</td>
<td></td>
<td>Likelihood of DMARD change, aOR (95% CI) G1: 1.09 (CI 0.57-2.12) G2: 0.56 (CI 0.29-1.06) G3: 0.30 (CI 0.12-0.79, p=0.014)</td>
<td></td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Risk of adverse drug reaction, aOR (95% CI) G1: 1.92 (CI 0.85-4.37) G2: 0.38 (CI 0.16-0.94, p&lt;0.05) G3: 0.33 (CI 0.08-1.38)</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Montecucco et al., 2012&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Patients met RA classification criteria, aged &gt;18 years, symptom duration &lt;12 mos, and had active disease according to the disease activity score.</td>
<td><strong>Interventions, dose:</strong>&lt;br&gt;G1: MTX 10 mg/wk (max 25 mg/wk) + PRED 12.5 mg/d for 2 wks&lt;br&gt;G2: MTX 10 mg/wk (max 25 mg/wk)</td>
<td><strong>Median disease duration, mos:</strong>&lt;br&gt;2.97-3.48</td>
<td><strong>At 12 months</strong>&lt;br&gt;DAS disease activity&lt;br&gt;Figure only data</td>
<td>Overall: NR</td>
<td>Medium</td>
</tr>
<tr>
<td>Country, Clinical Setting: Italy, University hospital clinic</td>
<td></td>
<td><strong>Baseline DAS, median:</strong>&lt;br&gt;5.0-5.2</td>
<td><strong>ACR20 response, %</strong>&lt;br&gt;NR</td>
<td>SAEs NR</td>
<td>Overall discontinuation&lt;br&gt;G1: 8.2&lt;br&gt;G2: 10.9</td>
<td></td>
</tr>
<tr>
<td>Study Design: RCT</td>
<td></td>
<td><strong>Baseline HAQ, median:</strong>&lt;br&gt;1.0-1.1</td>
<td><strong>LDA</strong>&lt;br&gt;G1: 80.2%&lt;br&gt;G2: 75.5%&lt;br&gt;p=0.44</td>
<td>Discontinuation because of AEs&lt;br&gt;G1: 5.5&lt;br&gt;G2: 9.1&lt;br&gt;p=0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N: 220</td>
<td></td>
<td><strong>MTX naive:</strong>&lt;br&gt;NR</td>
<td><strong>DAS remission, %</strong>&lt;br&gt;G1: 44.8&lt;br&gt;G2: 27.8&lt;br&gt;p=0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration: 1 year</td>
<td></td>
<td><strong>MTX inadequate responders:</strong>&lt;br&gt;NR</td>
<td><strong>SHS</strong>&lt;br&gt;NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Biologic non-responders:</strong>&lt;br&gt;NR</td>
<td><strong>HAQ</strong>&lt;br&gt;NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Seropositive (RF or CCP) (%):</strong>&lt;br&gt;NR</td>
<td><strong>SF-36</strong>&lt;br&gt;NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Baseline Sharp score, mean:</strong>&lt;br&gt;NR</td>
<td><strong>VAS pain</strong>&lt;br&gt;Figure only data&lt;br&gt;Mean difference: -8.8 (95% CI, -17.5 to -0.1); p=0.04</td>
<td></td>
<td>Specific AEs NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>At 9 months</strong>&lt;br&gt;VAS pain&lt;br&gt;Figure only data; p=NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>At 6 months</strong>&lt;br&gt;VAS pain&lt;br&gt;Figure only data; p=NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Montecucco et al., 2012</td>
<td>Adults (aged ≥ 18 yrs) fulfilling ACR criteria for RA with disease duration &lt; 3 yrs, active RA defined as ≥ 4 swollen joints and 4 tender joints (using a 28-joint count), rheumatoid factor or anti-CCP antibody positivity, ≥2 erosions on radiographs of hands/wrists/feet, prior CS use limited to ≤ 10 mg/day of PRED and stable ≥ 2 wsks prior to</td>
<td>Interventions, dose: G1 (immediate):  • MTX: Escalated to 20 mg/wk, or lower dose if no active tender/painful or swollen joints at wk 12 (oral)  • SSZ: 500 mg twice/day and, if tolerated, escalated to 1,000 mg twice/day  • HCQ: 200 mg twice/day  • Folic acid: 1 mg/day G2 (immediate):  • MTX: Escalated to 20 mg/wk, or lower dose if no active tender/painful or swollen joints at wk 12 (oral)  • ETN: 50 mg/wk (subcutaneous)  • Placebo  • Folic acid: 1 mg/day G3 (step-up):  • MTX: Escalated to 20 mg/wk, or lower dose if no active tender/painful or swollen joints at wk 12 (oral)</td>
<td>Erosive disease, %: NR</td>
<td>At 4 months VAS pain Figure only data Mean difference: -10.8 (95% CI, -19.1 to -2.5); p=0.01</td>
<td>Overall AEs: High</td>
<td></td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Moreland et al., 2012; O'Dell et al., 2013</td>
<td>Country, Clinical Setting: United States, Multicenter</td>
<td>Study Design: RCT</td>
<td>Overall N: 755</td>
<td>Study Duration: 2 yrs</td>
<td>Mean disease duration, mos: 2.9-4.5 Baseline DAS28-ESR among completers, mean: 5.8-5.9 Baseline modified HAQ among completers, mean: 1.0-1.1</td>
<td>DAS28-ESR disease activity, mean: G1: 2.9 (SD, 1.5) G2: 3.0 (SD, 1.4) G3: 2.8 (SD, 1.3) G4: 3.0 (SD, 1.4) DAS28-ESR disease activity, mean change from wk 48 (primary outcome): By arm: p=0.28 G1/2 vs. G3/4: p=0.55 G2/4 vs. G1/3: p=0.48 ACR20 response, %: Figure only data; p=NS ACR50 response, %: Figure only data; p=NS</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
</tbody>
</table>
| Author, yr, Study Name: Moreland et al., 2012[20], O’Dell et al., 2013[159] | screening, and no prior biologic therapy | - SSZ: 500 mg twice/day if DAS28-ESR ≥ 3.2 at wk 24; if tolerated, escalated to 1,000 mg twice/day (otherwise, placebo)  
- HCQ: 200 mg twice/day if DAS28-ESR ≥ 3.2 at wk 24 (otherwise, placebo)  
- Folic acid: 1 mg/day  
- MTX: Escalated to 20 mg/wk, or lower dose if no active tender/painful or swollen joints at wk 12 (oral)  
- ETN: 50 mg/wk (subcutaneous) if DAS28-ESR ≥ 3.2 at wk 24 (otherwise, placebo)  
- Placebo  
- Folic acid: 1 mg/day | Low-dose CS treatment at screening, %: 41.7  
Prior csDMARD use, %: 23.6  
MTX naive: 79.2  
MTX inadequate responders: NR  
Biologic non-responders, %: 0.5 (protocol exceptions)  
RF seropositive, %: 89.7  
RF negative/anti-CCP seropositive, %: 3.3  
Baseline mTSS among completers, mean: 4.1-6.5 | ACR70 response, %: Figure only data;  
DAS remission (< 2.6), %:  
G1: 59.1  
G2: 56.6  
G3: 56.5  
G4: 52.9  
p=0.93  
G1/2 vs. G3/4: p=0.36  
G2/4 vs. G1/3: p=0.43  
mTSS score, mean:  
G1: 7.3 (SD, 13.3)  
G2: 7.0 (SD, 16.6)  
G3: 6.2 (SD, 8.9)  
G4: 4.8 (SD, 7.2)  
Change in G1/2 vs. G3/4: p=0.74  
Change in G2/4 vs. G1/3: 0.64 vs. 1.69; p=0.047  
No radiographic progression (mTSS change < 0.5), %:  
G1: 64.9  
G2: 79.4  
G3: 68.3  
G4: 71.1  
p=0.33  
G1/2 vs. G3/4: p=0.56  
G2/4 vs. G1/3: p=0.02 | Discontinuation because of SAEs: G1/2: 2.7  
G3/4: 1.1  
Discontinuation because of lack of efficacy: G1/2: 3.7  
G3/4: 2.9  
Patient compliance: G1/2 vs. G3/4: p=0.74  
G2/4 vs. G1/3: p=0.76  
Specific AEs: NR |
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Moreland et al., 2012; O'Dell et al., 2013</td>
<td></td>
<td>Race, % white: 79.6</td>
<td>Erosive disease, %: NR</td>
<td>Modified HAQ, mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: 11.3</td>
<td></td>
<td>G1: 1.0 (SD, 0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Hispanic: 11.3</td>
<td></td>
<td>G2: 1.0 (SD, 0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 0.9 (SD, 0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 0.9 (SD, 0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At wk 48</td>
<td>No difference in HAQ functional capacity among groups (p=NR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At wk 24</td>
<td>(prior to initiating step-up)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28-ESR disease activity, mean change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/2: 3.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3/4: 4.2</td>
<td>p&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/2 &gt; G3/4: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/2 &gt; G3/4: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/2 &gt; G3/4: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Mottonen et al., 1999; Puolakka et al., 2004; Korpela et al., 2004; Makinen et al., 2007; Rantalaiho et al., 2010; Karstila et al., 2012; Rantalaiho et al., 2013; FIN-RACo</td>
<td>Adults (aged 18-65 yrs) fulfilling ACR criteria with symptom duration &lt; 2 yrs, active disease defined as ≥ 3 swollen joints and ≥ 3 of the following: ESR ≥ 28 mm/h, CRP &gt; 19 mg/L, morning stiffness ≥ 29 mins, &gt; 5 swollen joints and &gt; 10 tender joints; patients had no prior DMARD use and no glucocorticoid therapy within previous 2 wks</td>
<td>G1:</td>
<td>Mean disease duration, mos:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong></td>
<td>Finland, multicenter</td>
<td>MTX: Initiated at 7.5 mg/wk and increased to 10 mg/wk if patient did not achieve clinical improvement at 3 mos; could be tapered and then discontinued at 18 mos if remission achieved during first yr with initial combo</td>
<td>Baseline DAS, mean:</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HCQ: 300 mg/day</td>
<td>Baseline HAQ, mean:</td>
<td>G1: 0.9 (SD, 0.6) G2: 0.9 (SD, 0.6) Overall: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSZ: 500 mg/twice daily</td>
<td>MTX naïve, %:</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PNL: Initiated at 5 mg/day and increased to 7.5 mg/day if patient did not achieve clinical improvement at 3 mos; could be tapered and then discontinued at 9 mos if remission achieved during first yr with initial combo</td>
<td>MTX inadequate responders, %:</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2:</td>
<td>Biologic non-responders, %:</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSZ: Initiated at 2 g/day and increased to 3 g/day if clinically indicated at 3 mos</td>
<td>RF seropositive, %:</td>
<td>68.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Patients switched to 7.5-15 mg/wk MTX at 6 mos if an AE occurred or clinical response &lt; 25%</td>
<td>Baseline Sharp score, mean:</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Larsen score, median:</td>
<td>G1: 2 (IQR, 0-4) G2: 2 (IQR, 0-8) Overall: NR</td>
<td>Larsen score, median:</td>
<td>G1: 2 (IQR, 0-4) G2: 2 (IQR, 0-8) Overall: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 7.3 (range 2-22) G2: 8.6 (range 2-23)</td>
<td>Baseline DAS28, mean:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Baseline DAS28 disease activity:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %:</td>
<td>G1: 78 (95% CI, 69 to 80) G2: 84 (95% CI, 75 to 90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %:</td>
<td>G1: 71.1 G2: 58.1 p=0.058</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission, %:</td>
<td>G1: 68 G2: 41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sustained DAS28 remission, % (95% CI):</td>
<td>G1: 51 (95% CI 39 to 62) G2: 16 (95% CI 10 to 24) P &lt; 0.001 OR: 5.58 (95% CI 2.60-11.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR remission, %:</td>
<td>G1: 42 G2: 20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall AEs:</td>
<td>G1: 70.1 G2: 71.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SAEs:</td>
<td>G1: 3.1 G2: 5.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall discontinuation:</td>
<td>G1: 10.3 G2: 7.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to AEs:</td>
<td>G1: 23.7 G2: 22.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation due to lack of efficacy:</td>
<td>G1: 1.0 G2: 0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Patient adherence:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific AEs</td>
<td>AAT and AP &gt; 2x normal:</td>
<td>G1: 11.3 G2: 23.5 p=0.026</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Mottonen et al., 1999; Puolakka et al., 2004; Korpela et al., 2004; Makinen et al., 2007; Rantalaiho et al., 2010; Karstila et al., 2012; Rantalaiho et al., 2013; FIN-RACo (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sustained ACR remission, % (95% CI)</strong></td>
<td>G1: 14 (95% CI 7 to 23)</td>
<td>G2: 3 (95% CI 1 to 9)</td>
<td>P = 0.013</td>
<td>OR: 4.61 (95% CI 1.17-16.99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clinical remission, %:</strong></td>
<td>G1: 37.1</td>
<td>G2: 18.4</td>
<td>p=0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sharp score:</strong></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Larsen score median:</strong></td>
<td>G1: 4 (IQR, 0-14)</td>
<td>G2: 12 (IQR, 4-20)</td>
<td>p=0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Median increase in Larsen Score:</strong></td>
<td>G1: 1.5</td>
<td>G2: 2.0 (p&lt;0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Mottonen et al., 1999; Puolakka et al., 2004; Korpela et al., 2004; Makinen et al., 2007; Rantalaiho et al., 2010; Karstila et al., 2012; Rantalaiho et al., 2013</td>
<td>N: G1: 97 G2: 98</td>
<td>Radiographic evidence of erosions, %: 48.2</td>
<td>HAQ, mean change: G1: -0.6 (95% CI, -0.7 to -0.4) G2: -0.6 (95% CI, -0.8 to -0.5)</td>
<td>Cardiovascular Events: G1: 1 MI G2: 2 MIs</td>
<td></td>
</tr>
<tr>
<td>Study Design: RCT</td>
<td>Overall N: 199</td>
<td>Mean age, yrs: G1: 47 (range 23-65) G2: 48 (range 20-65) Overall: NR</td>
<td>Median work disability per pt-observation yr, days: G1: 12.4 G2: 32.2 (p=0.008)</td>
<td></td>
<td>Malignancies: 1 prostate cancer; 1 multiple myeloma</td>
<td></td>
</tr>
<tr>
<td>Study Duration: 2 yrs</td>
<td></td>
<td>Sex, % female: 62.1</td>
<td>At 1 yr ACR50 response, %: G1: 70.1 G2: 57.1 p=0.028</td>
<td></td>
<td>URTI: 1 pneumonia</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td>Clinical remission, %: G1: 24.7 G2: 11.2 p=0.011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: NR</td>
<td>DAS28 remission, % G1: Figure only (Fig. 2) G2: Figure only (Fig. 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td>Sustained DAS28 remission, % G1: 57.0 G2: 23.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR remission, % G1: Figure only (Fig. 2) G2: Figure only (Fig. 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sustained ACR remission, % G1: 16.5 G2: 3.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>At 6 mos ACR20 response, %: G1: 80 (95% CI, 71 to 88) G2: 78 (95% CI, 69 to 86)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics

<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr,</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Name:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mottonen et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999&lt;sup&gt;22&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Puolakka et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004&lt;sup&gt;102&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korpela et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004&lt;sup&gt;101&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Makinen et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007&lt;sup&gt;144&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rantalaiho et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010&lt;sup&gt;142&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karstila et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012&lt;sup&gt;145&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rantalaiho et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013&lt;sup&gt;143&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong> Nam et al., 2014</td>
<td>Patients aged 18-80, meeting ACR criteria for RA, with 3-12 mos symptom duration, active disease (DAS&gt;2.4) and DMARD naive</td>
<td>Interventions, dose: G1: MTX (10 mg/wk to max tolerated dose) + IFX (3 mg/kg) G2: MTX (10 mg/wk to max tolerated dose) + Intraocular Methyl-PNL (250 mg single dose) + Placebo</td>
<td>Median disease duration, mos: 1.2</td>
<td>At week 78 (Open Label)</td>
<td>Overall</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> IDEA</td>
<td>Overall N: 112</td>
<td>MTX: 10 mg/wk to 20 mg or max tolerated dose by wk 6</td>
<td>Baseline DAS, mean: 3.56-4.05</td>
<td>DAS disease activity</td>
<td>G1: 98.2</td>
<td>G2: 94.7</td>
</tr>
<tr>
<td><strong>Study Design:</strong> RCT</td>
<td>Study Duration: 78 wks (1-26 wks blinded, 26-78 wks open-label)</td>
<td>IFX: Max dose 1000 mg, delivered via infusion at wks 0, 2, 6, 14, 22</td>
<td>Baseline HAQ-DI, mean: 1.34-1.43</td>
<td>ACR20 response, %</td>
<td>G1: 36.4</td>
<td>G2: 15.8</td>
</tr>
<tr>
<td><strong>Overall N:</strong> 112</td>
<td>N: G1: 55 G2: 57</td>
<td>Methyl-PNL/Placebo: Delivered via infusion at wk 0; placebo delivered at wks 2, 6, 14, 22, 26, 38, 50, 68 and 78</td>
<td>MTX naive: 100</td>
<td>ACR50 response, %</td>
<td>G1: 20</td>
<td>G2: 24.6</td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 78 wks (1-26 wks blinded, 26-78 wks open-label)</td>
<td>Mean age, yrs: 52.9-53.7</td>
<td>Prior csDMARD use, %: 0</td>
<td>ACR70 response, %</td>
<td>G1: 10.5</td>
<td>G2: 1.8</td>
<td></td>
</tr>
<tr>
<td><strong>Mean age, yrs:</strong> 52.9-53.7</td>
<td>Sex, % female: 68.8</td>
<td>MTX inadequate responders: 100</td>
<td>DAS28 remission, %</td>
<td>G1: 54.3</td>
<td>G2: 65.3</td>
<td></td>
</tr>
<tr>
<td><strong>Sex, % female:</strong> 68.8</td>
<td>Race, % white: NR</td>
<td>Biologic non-responders: NR</td>
<td>DAS remission, %</td>
<td>G1: 47.7</td>
<td>G2: 50.0</td>
<td></td>
</tr>
<tr>
<td><strong>Race, % white:</strong> NR</td>
<td>Prior CS use, %: 0</td>
<td>Prior CS use, %: 0</td>
<td>mTSS total score, mean (SD)</td>
<td>G1: 1.69 (SD, 3.28)</td>
<td>G1: 5.5</td>
<td>G2: 1.8</td>
</tr>
<tr>
<td><strong>Prior CS use, %:</strong></td>
<td>0</td>
<td>RF seropositive (%): 55</td>
<td>Adjusted difference (95% CI): mTSS total score, mean (SD)</td>
<td>G2: 3.19 (SD, 7.75)</td>
<td>p=0.253</td>
<td>G2: 3.19 (SD, 7.75)</td>
</tr>
<tr>
<td><strong>RF seropositive (%):</strong></td>
<td>55</td>
<td>Erosion disease: NR</td>
<td>Mean change in HAQ-DI, mean (SD)</td>
<td>G1: -0.85 (SD, 0.60)</td>
<td>Adjusted difference (95% CI): mTSS total score, mean (SD)</td>
<td>G2: -0.79 (SD, 0.54)</td>
</tr>
<tr>
<td><strong>Mean change in HAQ-DI, mean (SD):</strong></td>
<td>0.60-9.23</td>
<td>SF-36</td>
<td><strong>SAEs</strong></td>
<td>G1: 36.4</td>
<td>G2: 15.8</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline mTSS score, mean:</strong> NR</td>
<td><strong>Overall discontinuation</strong></td>
<td>G1: 20</td>
<td>G2: 24.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Infection – pulmonary/upper respiratory</strong></td>
<td><strong>Patient adherence</strong></td>
<td>G1: 5.5</td>
<td>G2: 1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>Discontinuation because of AEs</strong></td>
<td>G1: 10.5</td>
<td>G2: 1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td><strong>Infection – pulmonary/upper respiratory</strong></td>
<td>G1: 5.5</td>
<td>G2: 1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Nam et al., 2014 IDEA (continued)</td>
<td></td>
<td></td>
<td></td>
<td>At week 50 (Open Label)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS disease activity</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR 20/50/70, %</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EULAR remission, %</td>
<td>G1: 16.5</td>
<td>G2: 19.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission, %</td>
<td>G1: 55.7</td>
<td>G2: 49.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mTSS total score, mean (SD)</td>
<td>G1: 1.20 (SD, 2.27)</td>
<td>G2: 2.81 (SD, 6.88)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p=0.132</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adjusted difference (95% CI):</td>
<td>−1.45 (CI −3.35 to 0.45)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>At week 26</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% achieving LDA score, DAS28 ≤3.2</td>
<td>G1: 64.4</td>
<td>G2: 66.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td>G1: 71.0</td>
<td>G2: 75.2</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Nam et al., 2014&lt;sup&gt;96&lt;/sup&gt; IDEA (continued)</td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 54.0</td>
<td>G2: 55.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 32.7</td>
<td>G2: 31.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remission, (DAS28 &lt;1.6), %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1: 40.6</td>
<td>G2: 50.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mTSS total score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 0.83 (SD, 1.69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 1.52 (SD, 4.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>p=0.291</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjusted difference (95% CI):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>~0.59 (CI -1.70 to 0.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean change in HAQ-DI, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: -0.70 (SD, 0.56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: -0.61 (SD, 0.47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF-36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At week 14 % achieving LDA score, DAS28 ≤3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 55.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 54.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Nam et al., 2014&lt;sup&gt;[9]&lt;/sup&gt; IDEA (continued)</td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td>NR</td>
<td>High</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Porter et al., 2016&lt;sup&gt;[8]&lt;/sup&gt; ORBIT</td>
<td>Patients were aged &gt;18, met 1987 ACR criteria for RA, and had a DAS28 score &gt;5.1. All had previously attempted treatment with ≥2 csDMARDs, were seropositive for RF or CCP, and were biological treatment naïve. All patients were not pregnant, breastfeeding, or of childbearing potential.</td>
<td>Interventions, dose: G1: RTX 1 g on days 1 and 15 with premedication 30 min before of methylprednisolone 100mg IV, paracetamol 1gram, chlorphenamine 10mg, and after 26 wks if patient responded to treatment but had persistent disease activity (DAS&gt;3.2). If flare (&gt;1.2 increase in DASESR), early retreatment &gt;20 weeks was allowed G2: TNF inhibitor – ADA 40 mg every other week subcutaneously, or ETN 50 mg/wk subcutaneously TNF inhibitor (either ADA or ETAN provided according to patient’s and rheumatologist’s choice</td>
<td>Mean disease duration, mos: G1: 8.0 G2: 6.7 Baseline DAS, mean: G1: 6.2 (0.9) G2: 6.2 (1.1) Baseline HAQ, mean: G1: 1.7 G2: 1.8 MTX naïve, %: 0 MTX inadequate responders, %: NR MTX intolerance, %: G1: 26 G2: 25</td>
<td>At 1 yr (primary outcome) DAS disease activity, mean change G1: -2.6 (SD, 1.4) G2: -2.4 (SD, 1.5) p=0.24 ACR20 response, % G1: 66 G2: 71 OR (95% CI)=0.8 (0.5-1.4) ACR50 response, % G1: 49 G2: 45 OR (95% CI)=1.2 (0.7-1.9) ACR70 response, % G1: 23 G2: 26 OR (95% CI)=0.8 (0.5-1.5)</td>
<td>Overall: G1: 95 G2: 95 SAEs G1: 25.7 G2: 17.2 Overall discontinuation G1: 18.8 G2: 17.7 Discontinuation because of AEs G1: 1.4 G2: 1.3 Patient adherence See comment</td>
<td></td>
</tr>
</tbody>
</table>
### Study Characteristics

**Author, yr, Study Name:** Porter et al., 2016

**ORB1T (continued)**

### Study Population Summary

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G1: 165</th>
<th>G2: 164</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs:</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>

### Interventions and Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G1: 165</th>
<th>G2: 164</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior csDMARD use, %:</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Biologic non-responders, %:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Seropositive (RF or CCP) (%):</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Baseline Sharp score, mean:</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Erosive disease, %:</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>

### Baseline Disease and Treatment Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G1: 23</th>
<th>G2: 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAS28 remission (DAS28 ESR &lt;2.6), %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHS</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>HAQ mean change from baseline</td>
<td>G1: -0.49 (SD, 0.6)</td>
<td>G2: -0.38 (SD 0.5)</td>
</tr>
<tr>
<td>EQ-5D mean change from baseline</td>
<td>G1: 0.2 (SD, 0.4)</td>
<td>G2: 0.3 (SD, 0.3)</td>
</tr>
</tbody>
</table>

### Health Outcomes

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G1: 61</th>
<th>G2: 65</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACR20 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)=0.8 (0.5-1.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>G1: 37</th>
<th>G2: 41</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACR50 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)=0.9 (0.5-1.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Adverse Events (%)

<table>
<thead>
<tr>
<th>Specific AEs</th>
<th>Infections:</th>
<th>Infections:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 53.5</td>
<td>G2: 70.9</td>
<td></td>
</tr>
<tr>
<td>Death:</td>
<td>G1: 1 (elbow prosthesis infection)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2: 1 (myocardial infarction)</td>
<td></td>
</tr>
</tbody>
</table>

**Specific AEs**

<table>
<thead>
<tr>
<th>Infections:</th>
<th>Infections:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: 53.5</td>
<td>G2: 70.9</td>
</tr>
<tr>
<td>Injection site reactions: p=0.003</td>
<td></td>
</tr>
</tbody>
</table>

Death:

<p>| G1: 1 (elbow prosthesis infection) |
| G2: 1 (myocardial infarction) |</p>
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Porter et al., 2016[^1] ORBIT (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 15</td>
<td>G2: 17</td>
<td>OR (95% CI)=0.8 (0.5-1.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 14</td>
<td>G2: 16</td>
<td>OR (95% CI)=0.9 (0.4-1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean change from baseline</td>
<td>G1: -0.44 (SD, 0.6) G2: -0.31 (SD, 0.6) p=0.0391</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EQ-5D mean change from baseline</td>
<td>G1: 0.2 (SD, 0.4) G2: 0.3 (SD, 0.4) p=0.9048</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Quinn et al., 2005</td>
<td>Patients with RA diagnosis meeting 1987 ACR criteria for RA with &lt;1 yr symptoms, no prior treatment with DMARDs or oral corticosteroids, MCP joint involvement, stable dosage of NSAIDs for 2 wks prior to screening, and poor prognosis according to PISA scoring system</td>
<td><strong>Interventions, dose:</strong> G1: • MTX: Beginning at 7.5 mg/wk, rapidly increased to 25 mg/wk in the presence of remaining synovitis • IFX: 3 mg/kg infusion at wks 0, 2, 6 and every 8 wks thereafter for 46 wks G2: • MTX: Beginning at 7.5 mg/wk, rapidly increased to 25 mg/wk in the presence of remaining synovitis</td>
<td><strong>Baseline DAS28, median:</strong> G1: 6.3 (IQR, 5.6-6.5) G2: 6.9 (IQR, 6.1-7.9)</td>
<td><strong>At 2 yrs (followup)</strong> ACR20 response, % G1: 70 G2: 50 ACR50 response, % G1: 70 G2: 50 ACR70 response, % G1: 67 G2: 30</td>
<td><strong>Overall AEs:</strong> Overall: 15 SAEs: NR</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong> RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall N:</strong> 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 2 yrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Mean disease duration, mos:** 6.0-7.4 mos 

**p<0.05**

**DAS28-4<2.6 remission, %** G1: 70 G2: 20

**SHS, mean change in total score from baseline** G1: 10 G2: 12

**Discontinuation because of lack of efficacy:** NR

**Patient adherence:** NR

**Specific AEs:** NR
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Quinn et al., 2005</td>
<td>N: G1: 10 G2: 10</td>
<td>Placebo</td>
<td>N: 52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs: 66.7</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, %: NR</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF seropositive, %: 65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Sharp score, mean: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erosive disease, %: Figure only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline HAQ, median (IQR): G1: 1.3 (IQR, 0.88) G2: 1.3 (IQR, 0.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX naïve, %: 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX inadequate responders, %: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologic non-responders: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**At 54 weeks**

- DAS28 disease activity score median change (IQR)
  - G1: Figure only (Fig 2)
  - G2: Figure only (Fig 2)

- ACR20 response, %
  - G1: 80
  - G2: 60

- ACR50 response, %
  - G1: 78
  - G2: 40
  - p<0.05

- ACR70 response, %
  - G1: 67
  - G2: 30
  - p<0.05

- DAS28-4<2.6 remission, %
  - G1: Figure only (Fig 6)
  - G2: Figure only (Fig 6)

- SHS
  - NR

- HAQ, % change in median functional score
  - Figure only, but significant functional benefit favoring G1>G2 (p=0.05)

- SF-36
  - NR

**At 14 weeks**

- DAS28 disease activity score median change (IQR)
  - G1: Figure only (Fig 2)
  - G2: Figure only (Fig 2)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, yr, Study Name: Quinn et al., 2005</td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %</td>
<td>G1: 60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %</td>
<td>G1: 60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %</td>
<td>G1: 60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 disease remission</td>
<td>G1: Figure only (Fig 2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only (Fig 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ, % change in median functional score</td>
<td>G1: Figure only (Fig 4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only (Fig 4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sharp score, mean change from baseline</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Author, year, study name, if applicable</td>
<td>Schipper et al., 2009&lt;sup&gt;36&lt;/sup&gt; Nijmegen RA Inception Cohort</td>
<td>Adults age ≥18 yrs with RA of &lt;1 yr duration diagnosed according to 1987 ACR revised criteria who had attempted SSZ treatment as first or second DMARD but were otherwise DMARD-naive</td>
<td>Comparisons (dosage and frequency) G1: MTX (7.5 mg/wk; max 30 mg/wk) G2: SSZ (750 mg/d; max 3 g/d) + MTX (7.5 - 30 mg/wk)</td>
<td>Median disease duration, wks: 14-47</td>
<td>At 52 weeks Overall discontinuation G1: 33.9 G2: 50 (P=0.013, mainly driven by events during first 6 months)</td>
<td>High</td>
</tr>
<tr>
<td>Country and setting</td>
<td>Netherlands, outpatient clinics</td>
<td></td>
<td>Baseline DAS28, mean 4.9-5.1</td>
<td>Baseline HAQ: NR</td>
<td>Discontinuation because of AEs G1: 18.5 G2: 11.3</td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>Observational (prospective cohort)</td>
<td>Mean age (years) 61.8-63.8</td>
<td>Prior CS use, %: 8-9</td>
<td></td>
<td>Discontinuation because of AEs G1: 14.5 G2: 8.5</td>
<td></td>
</tr>
<tr>
<td>Overall N</td>
<td>230</td>
<td>Sex, % female 70-74</td>
<td>Prior csDMARD use, %: Other than SSZ: 13-15</td>
<td>ACR 20/50/70, %: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of study</td>
<td>1 yr</td>
<td>Race, % white NR</td>
<td>MTX naïve, %: NR</td>
<td>EULAR good or moderate response, %: G1: 53 G2: 51 (P=NS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MTX inadequate responders: 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Biologic non-responders: 100 (to SSZ)</td>
<td>At 6 mos DAS28, mean difference in change from baseline (SD): G1: -0.9 (1.3) G2: -0.8 (1.3)</td>
<td>At 6 months Overall discontinuation G1: 18.5 G2: 31.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Seropositive (RF or CCP), %: RF(+): 73-81</td>
<td>Adjusted between-group difference (SE): -0.05 (0.16); P=0.737</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline Sharp score: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Erosive disease, %: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Soubrier et al., 2009; GUEPARD</td>
<td>Adults (aged ≥18 yrs) fulfilling ACR criteria for RA &lt; 6 mos, DAS28-ESR ≥ 5.1, no prior MTX or biologic use</td>
<td><strong>Interventions, dose:</strong>&lt;br&gt;<strong>G1:</strong>&lt;br&gt;• ADA: 40 mg every other wk; stopped at wk 12 if DAS28 &lt; 3.2; restarted for 12 wks if relapse occurred, and then increased to 40 mg/wk if DAS28 remained &gt; 3.2 after 12 wks and tapered then stopped if successful, otherwise ETN (25 mg twice/wk) initiated for 12 wks; ETN stopped if successful after 12 wks and restarted if relapse occurred; if ETN failed, LEF initiated for 12 wks&lt;br&gt;• MTX: initiated with 0.3 mg/kg/wk (adjusted to max 20 mg/wk); tapered to 7.5 mg/wk if DAS 28 &lt; 2.6 for ≥ 6 mos; initial dose reintroduced if disease activity flared up after tapering&lt;br&gt;<strong>G2:</strong>&lt;br&gt;• MTX: initiated with 0.3 mg/kg/wk (adjusted to max 20 mg/wk); tapered to 7.5 mg/wk if DAS 28 &lt; 2.6 for ≥ 6 mos; initial dose reintroduced if disease activity flared up after tapering; ADA (40 mg every other wk or 40 mg/wk), ETN (25 mg twice/wk), or LEF added if insufficient response at wk 12 or later</td>
<td><strong>Median disease duration, mos:</strong> 4.4&lt;br&gt;<strong>DAS28, mean:</strong> 6.2 (SD 0.8)&lt;br&gt;<strong>HAQ, mean:</strong> 1.4-1.7&lt;br&gt;<strong>MTX naïve, %:</strong> 100&lt;br&gt;<strong>Prior csDMARD use, %:</strong> 0&lt;br&gt;<strong>MTX inadequate responders, %:</strong> NA&lt;br&gt;<strong>Biologic non-responders, %:</strong> NA&lt;br&gt;<strong>Prior CS use, %:</strong> 15.4&lt;br&gt;<strong>RF seropositive, %:</strong> 73.8&lt;br&gt;<strong>anti-CCP seropositive, %:</strong> 73.1&lt;br&gt;<strong>Sharp score, mean:</strong> 2.4-7.5&lt;br&gt;<strong>Erosive disease, %:</strong> 34.4</td>
<td><strong>At 1 yr (change from wk 12)</strong>&lt;br&gt;<strong>DAS28-ESR disease activity, mean:</strong>&lt;br&gt;<strong>G1:</strong>&lt;br&gt;<strong>G2:</strong>&lt;br&gt;<strong>ACR20 response, %:</strong>&lt;br&gt;<strong>G1:</strong> 85&lt;br&gt;<strong>G2:</strong> 81&lt;br&gt;<strong>ACR50 response, %:</strong>&lt;br&gt;<strong>G1:</strong> 67&lt;br&gt;<strong>G2:</strong> 68&lt;br&gt;<strong>ACR70 response, %:</strong>&lt;br&gt;<strong>G1:</strong> 42&lt;br&gt;<strong>G2:</strong> 58&lt;br&gt;<strong>DAS remission, %:</strong>&lt;br&gt;<strong>G1:</strong> 39.4&lt;br&gt;<strong>G2:</strong> 59.4&lt;br&gt;<strong>P = 0.15</strong>&lt;br&gt;<strong>mTSS, mean change:</strong>&lt;br&gt;<strong>G1:</strong> 1.9 (SD 4) among 27&lt;br&gt;<strong>G2:</strong> 1.8 (SD 4.7) among 29&lt;br&gt;<strong>P = 0.18</strong>&lt;br&gt;<strong>HAQ, mean change:</strong>&lt;br&gt;<strong>G1:</strong> -1.02 (95% CI -1.24, -0.81)&lt;br&gt;<strong>G2:</strong> -0.93 (95% CI -1.17, -0.69)&lt;br&gt;<strong>P = 0.79</strong>&lt;br&gt;<strong>SF-36:</strong> Improvement in physical and mental components did not reach statistical significance (data NR)</td>
<td><strong>Overall AEs:</strong> NR&lt;br&gt;<strong>SAEs:</strong>&lt;br&gt;<strong>G1:</strong> 15.2&lt;br&gt;<strong>G2:</strong> 15.6&lt;br&gt;<strong>Overall discontinuation:</strong>&lt;br&gt;<strong>G1:</strong> 15.2&lt;br&gt;<strong>G2:</strong> 9.4&lt;br&gt;<strong>Discontinuation due to AEs:</strong> NR&lt;br&gt;<strong>Discontinuation due to lack of efficacy:</strong> NR&lt;br&gt;<strong>Patient adherence:</strong> NR&lt;br&gt;<strong>Specific AEs:</strong> NR</td>
<td>Medium (12 wks)</td>
</tr>
<tr>
<td>Country, Clinical Setting: France, multicenter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design: RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N: 65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration: 1 yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Soubrier et al., 2009; GUEPARD (continued)</td>
<td></td>
<td>Decision to adjust treatment made every 3 mos for patients not achieving DAS28 ≤ 3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean age, yrs:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 46.3 (SD 16.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 49.3 (SD 15.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sex, % female:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 80.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Pain (visual analogue scale):**
No difference (data NR)

**Fatigue (visual analogue scale):**
No difference (data NR)

**Patient global assessment (visual analogue scale):**
No difference (data NR)

**At 12 wks**

**ACR20 response, %:**
- G1: 84
- G2: 50
  Statistically significant (P = NR)

**ACR50 response, %:**
- G1: 66
- G2: 27
  Statistically significant (P = NR)

**ACR70 response, %:**
- G1: 44
- G2: 19
  Statistically significant (P = NR)

**DAS remission, %:**
- G1: 36.4
- G2: 12.5
  Statistically significant (P = NR)

**HAQ, mean change:**
- G1: -0.82 (95% CI: -1.11, -0.52)
- G2: -0.51 (95% CI: -0.72, -0.30)
  Statistically significant (P = 0.02)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr,</strong></td>
<td><strong>Studies</strong></td>
<td><strong>Interventions, dose:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td><strong>Study Name:</strong></td>
<td><strong>G1:</strong> MTX (20 mg/wk) + placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Clair et al.,</td>
<td>GUEPARD</td>
<td><strong>G2:</strong> MTX + IFX (3 mg/kg/wk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004;17</td>
<td></td>
<td><strong>G3:</strong> MTX + IFX (6 mg/kg/wk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al.,</td>
<td>ASPIRE</td>
<td><strong>N:</strong> G1: 298</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006;107</td>
<td></td>
<td>G2: 373</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al.,</td>
<td></td>
<td>G3: 378</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009;106</td>
<td></td>
<td><strong>Mean age, yrs:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janssen Research and</td>
<td><strong>Race, % white:</strong></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development, 2017;157</td>
<td></td>
<td>71.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPIRE</td>
<td></td>
<td><strong>Sex, % female:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Setting:</td>
<td><strong>Seropositive (RF or CCP) (%):</strong></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multinational,</td>
<td><strong>Baseline Sharp score, mean:</strong></td>
<td>71.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>university hospitals</td>
<td>11.2-11.6</td>
<td><strong>Discontinuation because of AEs:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design:**</td>
<td><strong>Discontinuation due to lack of efficacy:</strong></td>
<td>25.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td><strong>G1:</strong> 3.2</td>
<td><strong>G1:</strong> 9.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N:**</td>
<td><strong>G2:</strong> 9.5</td>
<td><strong>G2:</strong> 1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1049</td>
<td><strong>G3:</strong> 9.6</td>
<td><strong>G3:</strong> 3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Duration:**</td>
<td></td>
<td><strong>Overall:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54 wks</td>
<td><strong>Overall:</strong></td>
<td><strong>Serious Infections (≥1 infection):</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G1:</strong> 2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G2:</strong> 5.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G3:</strong> 5.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>p=0.02</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Overall:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Discontinuation:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G1:</strong> 25.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G2:</strong> 21.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G3:</strong> 23.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Discontinuation due to lack of efficacy:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G1:</strong> 9.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G2:</strong> 1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>G3:</strong> 3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>----------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td>Erosive disease, %: 80-84</td>
<td>mTSS score change: G1: 3.7, G2: 0.4, G3: 0.5 (G1 vs. G2, G3: p&lt;0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Clair et al., 2004;17</td>
<td></td>
<td></td>
<td>Changes in TSS by disease activity (remission, low, moderate, high) G1: 1.1, 2.2**, 3.9**, 5.8** G2: -0.2, -0.4, 0.6, 2.1. [COMPARED WITH G2: *p=0.05, **p=0.01]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al., 2006;197</td>
<td></td>
<td></td>
<td>HAQ &gt; 0.22, %: G1: 65.2, G2: 76.0, G3: 75.5 (G2 vs. G1; p=0.003) (G3 vs. G1; p&lt;0.004)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al., 2009;106</td>
<td></td>
<td></td>
<td>SF-36 PCS scores G1: 10.1, G2: 11.7, G3: 13.2 G3 vs. G1, p=0.003 G3 vs. G2, p=0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janssen Research and Development, 2017157</td>
<td></td>
<td></td>
<td>Employability: IFX + MTX (OR 2.4 [95% CI 2.23 to 2.61], p&lt;0.001) MTX (p=0.56) Combo has higher probability of improvement than MTX alone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPIRE (continued)</td>
<td></td>
<td></td>
<td>Net increase in employability, %: MTX + IFX: 8 MTX-only: 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Patient adherence NR Infusion or injection reaction: G1: 7 G2: 21 G3: 15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TB: G1: 0 G2: 0.8 G3: 0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Clair et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004;^{17}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006;^{107}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smolen et al.,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009;^{106}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janssen Research and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development, 2017;^{157}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPIRE (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Net change in actual employment, %:**
  - MTX + IFX: -0.5
  - MTX-only: -1.3 (p=NS)

- **Employability status changed from employable to unemployable, %:**
  - IFX: 8
  - MTX-only: 14 (p=0.05)

- **At weeks 30 to 54**
  - **HAQ:**
    - G1: 0.68
    - G2: 0.80
    - G3: 0.88;
    - (G2 vs. G1; p=0.03)
    - (G3 vs. G1; p<0.001)
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr:</strong></td>
<td>Svensson et al., 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High</td>
</tr>
</tbody>
</table>
| Country, Setting      | BARFOT Study #1 (1992-1995) | Patients with active RA diagnosed according to 1987 ACR revised criteria who were DMARD and glucocorticoid-naive. Patients not included if not seen within 1 yr of symptom or sign of RA | **Interventions, dose:**
|                       |                          |                                          |                                               |                |                   |           |
|                       |                          |                                          | G1: PRED 7.5-15 mg/d for 1-3 months + MTX (if needed) 5-15 mg/wk, dosages NR | Mean disease duration, mos: 6 mos |                   |           |
|                       |                          |                                          | G2: SSZ 2-3 g/day + PRED (if needed) up to 10 mg/d | Prior csDMARD use, %: 0 |                   |           |
|                       |                          |                                          | N: G1: 113 G2: 108 | Prior CS use, %: 0 |                   |           |
|                       |                          |                                          | Median age, yrs: 54 | MTX naive, %: 100 |                   |           |
|                       |                          |                                          | Sex, % female: 63 | RF seropositive, %: 56 (between-group difference, p=0.0005) | Remission, DAS28 <2.6, %
| Study Duration:       | 2 yrs                    |                                          | Race, % white: NR |                           | G1: 29 G2: 19 (p=0.095) |           |
| Overall N:            | 245                      |                                          |                                               | Larsen score, mean change from baseline |
|                       |                          |                                          |                                               | G1: 6.2 (SD, 12.2) G2: 4.1 (SD, 10.9, p=0.298) | HAQ mean change from baseline
|                       |                          |                                          |                                               |                            | G1: -0.35 (SD, 0.61) G2: -0.38 (SD, 0.55, p=0.752) |           |
|                       |                          |                                          |                                               | SF-36 outcome |                   |           |
|                       |                          |                                          |                                               | NR |                   |           |
| Study Design:         | RCT                      |                                          |                                               | At 3 months |                   |           |
|                       |                          |                                          |                                               | Figure only (Figure 2) |                   |           |

**Interventions, dose:**

- G1: PRED 7.5-15 mg/d for 1-3 months + MTX (if needed) 5-15 mg/wk, dosages NR
- G2: SSZ 2-3 g/day + PRED (if needed) up to 10 mg/d

**Baseline Disease and Treatment Characteristics:**

- Mean disease duration, mos: 6 mos
- Prior csDMARD use, %: 0
- Prior CS use, %: 0
- MTX naive, %: 100
- RF seropositive, %: 56 (between-group difference, p=0.0005)
- Baseline DAS, mean: 4.9-5.0
- DAS score >3.2, %: 92
- HAQ, median score: 0.9
- Larsen score, median: 4.0

**Health Outcomes:**

- At 2 yrs:
  - DAS disease activity NR
  - Good EULAR response, %
    - G1: 30
    - G2: 33
  - Moderate EULAR response, %
    - G1: 40
    - G2: 30
  - No EULAR response, %
    - G1: 30
    - G2: 37
  - Remission, DAS28 <2.6, %
    - G1: 29
    - G2: 19 (p=0.095)
  - Larsen score, mean change from baseline
    - G1: 6.2 (SD, 12.2)
    - G2: 4.1 (SD, 10.9, p=0.298)
  - HAQ mean change from baseline
    - G1: -0.35 (SD, 0.61)
    - G2: -0.38 (SD, 0.55, p=0.752)
  - SF-36 outcome
    - NR

**Adverse Events (%):**

- Overall: NR
- SAEs: NR
- Overall discontinuation
  - G1: 19.5
  - G2: 47.2
- Discontinuation because of AEs
  - G1: 11.5
  - G2: 33.3

**Patient adherence:**

Patients who stayed on the allocated treatment for 2 yrs called “completers.” Overall, one-third of patients were non-completers (19% from G1 and 47% from G2)

**Specific AEs:**

- NR
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson et al., 2005</td>
<td>Adults ages 18 to 80 yrs</td>
<td><strong>Interventions, dose:</strong> G1: PNL (7.5 mg/d) + DMARD (SSZ 2 g/day, or MTX mean dose 10 mg/wk)</td>
<td>Mean disease duration, mos: 5.8-6.5</td>
<td>At 4 years (followup) DAS disease activity NR</td>
<td>NR</td>
<td>Medium (High for 4 year outcomes)</td>
</tr>
<tr>
<td>Hafstrom et al., 2009</td>
<td>with active RA of ≤ 1 yr</td>
<td>G1a: Subset of G1 (PNL + DMARD) participants who agreed to participate in 4 year followup</td>
<td>DMARD naive, %: 100</td>
<td>ACR20/50/70 or EULAR response, % NR</td>
<td>Overall: NR</td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al., 2014</td>
<td>duration diagnosed according to 1987 ACR revised criteria who were DMARD and glucocorticoid-naive</td>
<td>G1b: Subset of G1a participants in remission</td>
<td>Corticosteroid naive, %: 100</td>
<td>DAS remission According to longitudinal analysis investigating the relationship between DAS remission and radiographic damage in patients randomized to G1a and G21: DAS remission during followup=10.5 (Wald x^2), p&lt;0.001</td>
<td>SAEs NR</td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al., 2014</td>
<td>Sweden, multicenter (6 centers)</td>
<td>G1c: Subset of G1a participants not in remission</td>
<td>MTX naive, %: 100</td>
<td></td>
<td>Overall discontinuation G1: 11.8 G2: 19.8</td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td>Adults ages 18 to 80 yrs</td>
<td>G1d: Subset of G1 (PNL + DMARD) who had radiographs of hands and feet at baseline and 2 yr followup</td>
<td>Baseline DAS, mean: 5.3-5.4</td>
<td></td>
<td>Discontinuation because of AEs G1: 1.7 G2: 0</td>
<td></td>
</tr>
<tr>
<td><strong>Country, Setting:</strong></td>
<td>Excluded for previous fragility fractures, pts &lt; 65 yrs</td>
<td></td>
<td>HAQ: 0.98-1.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson et al., 2005</td>
<td></td>
<td>G1e: Subset of G1 (PNL + DMARD), only including patients who had no history of prior CV events</td>
<td>RF Seropositive, %: 66</td>
<td>mTSS, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al., 2014</td>
<td></td>
<td>G2: DMARD only (SSZ 2 g/day, or MTX mean dose 11 mg/wk)</td>
<td>Baseline Sharp score, mean: 4.1-4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al., 2014</td>
<td></td>
<td>G2a: Subset of G2 (DMARD only) participants who agreed to participate in 4 year followup</td>
<td>Erosion score at baseline, mean: 1.9</td>
<td>mTSS change from baseline, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td>(continued)</td>
<td>G2b: Subset of G1a participants not in remission</td>
<td></td>
<td>G1a: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N:</td>
<td>259</td>
<td>G2c: Subset of G1a participants not in remission</td>
<td></td>
<td>G1b: 4.5 (IQR, 2.0-7.5)</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Study Duration:</td>
<td>2 yrs (4 and 10 yr followup)</td>
<td>G2d: Subset of G2 (DMARD only) who had radiographs of hands and feet at baseline and 2 yr followup</td>
<td></td>
<td>G1c: 12.0 (IQR, 4.0-24.5)</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2e: Subset of G2 (DMARD only), only including patients who had no history of prior CV events</td>
<td></td>
<td>G2a: NR</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>N:</td>
<td>G1: 119 (a: 64, b: 35, c: 29, d: 108, e: 112)</td>
<td></td>
<td></td>
<td>G1b: 6.5 (IQR, 1.5-12.0)</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2: 131 (a: 86, b: 26, c: 60, d: 117, e: 111)</td>
<td></td>
<td></td>
<td>G2c: 10.5 (IQR, 1.0-20.0)</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1b vs. G1c: p=0.006</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2b vs. G2c: p=0.466</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ score improvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: NR</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1b: NR</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1c: NR</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: NR</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2b: NR</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2c: NR</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a vs. G2a: p=0.034</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Patient adherence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specific AEs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nausea</td>
<td>G1: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.8</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Leukopenia</td>
<td>G1: 0.8</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 2.3</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rash</td>
<td>G1: 5.0</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 6.9</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 10 yrs (followup)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total incident CV event, %</td>
<td>15.2</td>
<td>13.5 (p=0.72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Incident ischaemic coronary event, %</td>
<td>6.2</td>
<td>9.0 (p=0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Incident ischaemic cerebrovascular event, %</td>
<td>8.9</td>
<td>4.5 (p=0.19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Death, %</td>
<td>G1: 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 8 (p=0.98)</td>
<td>8 (p=0.98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td>Mean age, yrs:</td>
<td>At 3 years (followup)</td>
<td></td>
<td>Risk of CV-related</td>
<td></td>
</tr>
<tr>
<td>Svensson et al., 2005;</td>
<td></td>
<td>51-59</td>
<td>DAS disease activity</td>
<td>NR</td>
<td>death in patients</td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al., 2014;</td>
<td></td>
<td>64</td>
<td>ACR20/50/70 response, %</td>
<td>NR</td>
<td>with DAS remission</td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al., 2014;</td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td>NR</td>
<td>compared with those</td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td>(continued)</td>
<td></td>
<td>mTSS, mean (SE)</td>
<td>NR</td>
<td>not in remission,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1a: Figure only</td>
<td></td>
<td>HR (95% CI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1b: NR</td>
<td></td>
<td>G1e: 0.30 (CI 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G1c: NR</td>
<td></td>
<td>to 1.1, p=0.087)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2a: Figure only</td>
<td></td>
<td>G2e: 0.42 (CI 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2b: NR</td>
<td></td>
<td>to 2.03, p=0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G2c: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HAQ</td>
<td>NR</td>
<td>Risk of CV-related</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>NR</td>
<td>death in patients</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with good EULAR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>response compared</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with those without</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>good response, HR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(95% CI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1e: 0.45 (CI 0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>to 1.70, p=0.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2e: 0.28 (CI 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>to 1.13, p=0.074)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Author, yr:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson et al., 2005;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al., 2014;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al., 2014;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DAS28 &lt; 2.6 disease remission, % achieved</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 55.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1a: 55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 32.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2a: 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1 vs. G2: p=0.0005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1a vs. G2a: p=0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HAQ mean score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: Figure only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: Figure only (p=0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HAQ, mean decrease from baseline:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.5 (SD, 0.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0.7 (SD, 0.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change from baseline in mTSS, median (IQR)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1d: 1.8 (IQR, 0.5-6.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2d: 3.5 (IQR, 0.5-10.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(p=0.019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change from baseline in mTSS, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1d: 5.2 (SD, 9.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2d: 9.1 (SD, 14.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SF-36</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 18 mos:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28 score, mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: Figure only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: Figure only (p=0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td>ACR20/50/70 or EULAR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 &lt; 2.6 disease remission, % achieved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: 53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: 34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a vs. G2a: p=0.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only (p=0.0005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHS outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At 1 yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 score, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2.7 (SD, 1.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 3.3 (SD, 1.5, p=0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20/50/70 or EULAR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 &lt; 2.6 disease remission, % achieved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 51.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: 49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 39.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: 42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2: p=0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a vs. G2a: p=0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td>HAQ mean score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005;[38]</td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only (p=0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al.,</td>
<td></td>
<td></td>
<td></td>
<td>HAQ, mean decrease from baseline:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014;[40]</td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.4 (SD, 0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al.,</td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.6 (SD, 0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014;[38]</td>
<td></td>
<td></td>
<td></td>
<td>mTSS change from baseline, median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td>Continued</td>
<td></td>
<td></td>
<td>G1d: 1.0 (IQR, 0-3.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2d: 2.0 (IQR, 0-5.0) (p=0.035)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mTSS change from baseline, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1d: 2.4 (SD, 4.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2d: 5.3 (SD, 9.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 6 months DAS28 score, mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: Figure only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: Figure only (p=0.0005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20/50/70 or EULAR</td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 &lt; 2.6 disease remission, % achieved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a: 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2a: 22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1a vs. G2a: P =0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson et al., 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ajeganova et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hafstrom et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BARFOT Study #2 (1995-1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)

<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>At 3 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DAS28 score, mean**
- G1: Figure only
- G2: Figure only (p=0.0005)

**ACR20/50/70 or EULAR**
- NR

**DAS28 < 2.6 disease remission, % achieved**
- G1: NR
- G1a: 35
- G2: NR
- G2a: 9
- G1a vs. G2a: p=0.0005

**HAQ mean score**
- G1: Figure only
- G2: Figure only (p=0.0005)

**SHS outcome**
- NR

**SF-36**
- NR
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Adults (aged 18 to 80 years) fulfilling ACR criteria for RA with disease duration between 8 wks and 4 yrs, active disease defined as swollen and tender joint counts ≥ 8 each and CRP ≥ 1 mg/dl, radiographic evidence of erosive damage attributable to RA if rheumatoid factor negative, and no prior MTX use</td>
<td>Interventions, dose:</td>
<td>Mean disease duration, yrs:</td>
<td>At wk 104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2011;30</td>
<td></td>
<td>G1: • MTX: 7.5 mg/wk escalated up to 20 m/wk by wk 8 (oral)</td>
<td>0.91-0.99</td>
<td>G1: 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigby et al., 2011;131</td>
<td></td>
<td>• RIT: 1,000 mg on days 1 and 15 (intravenous; infusions premedicated with 100 mg methylprednisolone)</td>
<td>Baseline DAS28 ESR, mean: 7.0-7.1</td>
<td>G2: 45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012;132</td>
<td></td>
<td>G2: • MTX: 7.5 mg/wk escalated up to 20 m/wk by wk 8 (oral)</td>
<td>Baseline HAQ, mean: 1.7-1.8</td>
<td>G3: 25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td></td>
<td>• RIT: 500 mg on days 1 and 15 (intravenous; infusions premedicated with 100 mg methylprednisolone)</td>
<td>Concomitant CS, %: 46.4 (of 748)</td>
<td>G1/2 vs. G3: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multinational, multicenter:</td>
<td></td>
<td>G3: • MTX: 7.5 mg/wk escalated up to 20 m/wk by wk 8 (oral)</td>
<td>Concomitant CS, %: 46.4 (of 748)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Design:</strong> RCT</td>
<td></td>
<td>• Placebo</td>
<td>Concomitant glucocorticoids (≤ 10 mg/day PNL or equivalent) and non-steroidal anti-inflammatory drugs were allowed with stable doses while intravenous or - muscular glucocorticoids and additional DMARDs were not allowed; repeat courses were permitted from wk 24 for patients with DAS28 ESR ≥ 2.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N: 755</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 2 yrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Tak et al., 2011;</td>
<td><strong>N:</strong></td>
<td>RF seropositive, %: 86.4 (of 748)</td>
<td>HAQ-DI response (decrease ≥0.22), %:</td>
<td>Specific AEs Infusion-related reaction:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rigby et al., 2011;</td>
<td>G1: 251</td>
<td></td>
<td>G1: 86</td>
<td>G1: 18.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tak et al., 2012;</td>
<td>G2: 252</td>
<td></td>
<td>G2: 84</td>
<td>G2: 14.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IMAGE (continued)</td>
<td>G3: 252</td>
<td></td>
<td>G3: 77</td>
<td>G3: 12.4</td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs:</td>
<td>G1: 47.9 (SD, 13.3)</td>
<td><strong>Baseline Genant-modified Sharp score, mean:</strong></td>
<td>G1 vs. G3: p&lt;0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2: 47.9 (SD, 13.4)</td>
<td>6.9-7.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G3: 48.1 (SD, 12.7)</td>
<td>Overall: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td>81.1 (of 748)</td>
<td>Erosive disease, %: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % black:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity, % Latino:</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G1: 47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigby et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G2: 42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G3: 25</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IMAGE (continued)</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1/2 vs. G3: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigby et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>DAS28 ESR remission, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G1: 31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G2: 25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigby et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td>G3: 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td>G2 vs. G3: p&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Genant-modified Sharp score</strong></td>
<td></td>
<td></td>
<td></td>
<td>Total score, mean change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HAQ-DI</strong></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.359</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean change:</strong></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.646</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 1.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3: p&lt;0.001 No radiographic progression (change ≤ 0), %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 58</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G3: p&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No radiographic progression</strong></td>
<td></td>
<td></td>
<td></td>
<td>HAQ-DI Mean change:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.916</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.905</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: -0.628</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/2 vs. G3: p&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Tak et al., 2011;</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 Mental component, mean change: G1: 6.662 G2: 6.181 G3: 4.848</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak et al., 2012;</td>
<td></td>
<td></td>
<td></td>
<td>SF-36 Mental component, mean change: G1: 6.662 G2: 6.181 G3: 4.848</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Tak et al., 2011; Rigby et al., 2011; Tak et al., 2012; IMAGE</td>
<td>Adults (aged ≥ 20 yrs) fulfilling ACR criteria for RA with disease duration ≤ 2 yrs, tender joint count ≥ 10, swollen joint count ≥ 8, CRP level ≥ 1.5 mg/dl or ESR ≥ 28 mm/h, and ≥ 1 joint erosion or rheumatoid factor positivity; no prior treatment</td>
<td>Interventions, dose: G1: • MTX: 6 mg/wk and increased to 8 mg/wk if ≥ 20% decrease in tender or swollen joint counts not achieved on/after wk 8 (oral) • ADA: 40 mg every other wk (subcutaneous) • Folic acid: 5 mg/wk G2: • MTX: 6 mg/wk and increased to 8 mg/wk if ≥ 20% decrease in tender or swollen joint counts not achieved on/after wk 8 (oral) • Placebo • Folic acid: 5 mg/wk</td>
<td>Mean disease duration, yrs: 0.3</td>
<td>At wk 24 DAS28-ESR disease activity: NR</td>
<td>Genant-modified Sharp score Total score, mean change: G1: 0.328 G2: 0.580 G3: 0.701 G1 vs. G3: p&lt;0.05</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Takeuchi et al., 2014; Yamanaka et al., 2014 HOPEFUL 1</td>
<td>Country, Clinical Setting: Japan Study Design: RCT Overall N: 334 Study Duration: 26 wks (with 6 month open label)</td>
<td>Baseline DAS28-ESR, mean: 6.6 Baseline DAS28 (CRP), mean: 5.8-5.9 Baseline HAQ-DI, mean: 1.1-1.3</td>
<td>At 26 wks DAS28-ESR disease activity, change in mean: G1: -2.9 G2: -1.7 DAS28 (CRP) disease activity, change in mean: G1: -2.9 G2: -1.7 ACR20 response, %: G1: 75.4 G2: 56.4 ACR50 response, %: G1: 64.3 G2: 38.7 ACR70 response, %: G1: 47.4 G2: 22.7</td>
<td>Overall AEs: Medium G1: 80.7 (376 events) G2: 71.8 (302 events) SAEs: G1: 0.6 G2: 0.6 Overall discontinuation: G1: 15.2 G2: 22.1 Discontinuation because of AEs: G1: 4.1 G2: 2.5 Moved to rescue: G1: 8.2 G2: 17.2 Patient adherence: NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td></td>
<td>Those experiencing &gt; 20% increase in tender and swollen joint counts at wks 12, 16, or 20 were eligible for open-label rescue treatment with 40 mg ADA every other week; those completing the 26 wk double-blind period were eligible for open-label ADA + MTX for an additional 26 wks</td>
<td>Prior CS use, %: 32.0</td>
<td>DAS28-ESR remission (&lt; 2.6), %:</td>
<td>Specific AEs: Injection-site reaction</td>
<td></td>
</tr>
<tr>
<td>Takeuchi et al., 2014</td>
<td></td>
<td></td>
<td>Prior csDMARD use, %: 48.2</td>
<td>G1: 31.0</td>
<td>G1: 10.5</td>
<td></td>
</tr>
<tr>
<td>Yamanaka et al., 2014</td>
<td></td>
<td></td>
<td>MTX naïve, %: 100</td>
<td>G2: 14.7</td>
<td>G2: 3.7</td>
<td></td>
</tr>
<tr>
<td>HOPEFUL 1 (continued)</td>
<td></td>
<td></td>
<td>MTX inadequate responders, %: NR</td>
<td>p&lt;0.001</td>
<td>p=0.02</td>
<td></td>
</tr>
<tr>
<td>N: G1: 171 G2: 163</td>
<td></td>
<td></td>
<td>Biologic non-responders, %: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs: G1: 54.0 (SD, 13.1) G2: 54.0 (SD, 13.2)</td>
<td></td>
<td></td>
<td>RF seropositive, %: 84.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall: NR</td>
<td></td>
<td></td>
<td>anti-CCP seropositive, %: 84.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female: 81.4</td>
<td></td>
<td></td>
<td>Baseline mTSS score, mean: G1: 13.6 (SD, 22.3) G2: 13.6 (SD, 17.4) Overall: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white: NR</td>
<td></td>
<td></td>
<td>Erosive disease, %: NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % black: NR</td>
<td></td>
<td></td>
<td>No radiographic progression, %: G1: 62.0 G2: 35.4 (of 161) p&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity, % Latino: NR</td>
<td></td>
<td></td>
<td>Association of LDA at baseline with no radiographic progression (subgroup analysis – multivariate regression) G1: Not associated (p=NS) G2: Significantly associated (p=0.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Takeuchi et al., 2014; Yamanaka et al., 2014</td>
<td>HOPEFUL 1 (continued)</td>
<td></td>
<td>HAQ-DI:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -0.6 (SD, 0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -0.4 (SD, 0.6)</td>
<td></td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Response (&lt; 0.5), %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 60.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 36.8</td>
<td></td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 20 wks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28-ESR disease activity, change in mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 (CRP) disease activity, change in mean:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: -2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: -1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 78.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 62.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 62.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 37.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 36.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 16.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Takeuchi et al., 2014; Yamanaka et al., 2014</td>
<td>HOPEFUL 1 (continued)</td>
<td>At 16 wks</td>
<td>DAS28-ESR disease activity, change in mean: G1: -2.6 G2: -1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DAS28 (CRP) disease activity, change in mean: G1: -2.6 G2: -1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR20 response, %: G1: 74.8 G2: 54.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %: G1: 59.6 G2: 31.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %: G1: 31.0 G2: 14.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr,</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Study Name:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takeuchi et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yamanaka et al., 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HOPEFUL 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)

<table>
<thead>
<tr>
<th></th>
<th>At 12 wks</th>
<th>DAS28-ESR disease activity, change in mean:</th>
<th>G1: -2.5</th>
<th>G2: -1.4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DAS28 (CRP) disease activity, change in mean:</td>
<td>G1: -2.5</td>
<td>G2: -1.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACR20 response, %:</td>
<td>G1: 76.6</td>
<td>G2: 54.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACR50 response, %:</td>
<td>G1: 53.2</td>
<td>G2: 26.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACR70 response, %:</td>
<td>G1: 25.7</td>
<td>G2: 8.0</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Todoerti et al., 2010&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Patients meeting ACR criteria for RA with symptom duration &lt;12 mos</td>
<td><strong>Interventions, dose:</strong> G1: Low-dose oral PRED + MTX G2: MTX only Both treatments were DAS driven step-up protocols MTX: 10 mg/wk; Increased to 15 mg/wk and then to 20 mg/wk if LDA (DAS ≤ 2.4) not reached during followup visits Low-dose PRED: 12.5 mg/d for wks 1-2 then 6.25 mg/d</td>
<td><strong>Median disease duration, mos (IQR):</strong> 3 (1.93-5.4) <strong>Baseline DAS, mean:</strong> 3.74 (SD, 0.88) <strong>Baseline HAQ, median:</strong> 1.19 (IQR, 0.63-1.88) MTX naive: NR MTX inadequate responders: NR Biologic non-responders: NR Seropositive (RF or CCP) (%): RF+: 41.9-46.7 CCP+: 28.6-29.7 Baseline Sharp score, mean: NR Erosive disease, %: NR</td>
<td><strong>DAS remission, %</strong> At 18 mos G1: 76.7 G2: 33.3 (p=0.01) OR (95% CI) for probability of still being in remission over first 6 mos after first year of txmt: 4.480 (1.35-14.82) (p=0.014) P based on GEE analysis &lt;0.001 At 1 yr G1: 39.7 G2: 30.6 (p=0.290) OR (95% CI): 1.965 (1.214 to 3.182) (p=0.006) for probability of being in remission within 1 yr DAS more suppressed in G1 than G2 (P &lt; 0.001, based on GEE analysis) At 9 mos G1: 35.2 G2: 25.9 (p=0.239) At 6 mos G1: 26.3 G2: 16 (p=0.082) At 4 mos G1: 25.5 G2: 8 (p=0.001) At 2 mos G1: 14.9 G2: 7 (p=0.112)</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Author, year, study name, if applicable</td>
<td>Adults (aged ≥ 18 yrs) fulfilling ACR criteria for RA with symptom duration &lt; 1 yr, DAS28 &gt; 3.2, no prior DMARD use, no oral glucocorticoid or stable glucocorticoid therapy for ≥ 4 wks of ≤ 10 mg/day PRED (or equivalent), and for whom MTX ≤ 20 mg/wk had not lowered their DAS28 to ≤ 3.2 during the first 3 mos of disease treatment</td>
<td>Comparisons (dosage and frequency)</td>
<td>Mean disease duration, mos: 6.2-6.3</td>
<td>At 2 yrs</td>
</tr>
<tr>
<td>van Vollenhoven et al., 2009</td>
<td></td>
<td>G1:</td>
<td>ACR20 response %</td>
<td>G1: 33</td>
</tr>
<tr>
<td>Eriksson et al., 2013</td>
<td></td>
<td>G2:</td>
<td>(p=0.259)</td>
<td>G2: 40</td>
</tr>
<tr>
<td>van Vollenhoven et al., 2012</td>
<td></td>
<td>G1b: Obese (BMI ≥ 30) subpopulation of G1</td>
<td>ACR50 response%</td>
<td>G1: 22</td>
</tr>
<tr>
<td>Rezaei et al., 2013</td>
<td></td>
<td>G1c: Normal (BMI &lt;25) subpopulation of G1</td>
<td>ACR70 Response %</td>
<td>G1: 14</td>
</tr>
<tr>
<td>Eriksson et al., 2016</td>
<td></td>
<td>G2:</td>
<td></td>
<td>G2: 16</td>
</tr>
<tr>
<td>Levitsky et al., 2015</td>
<td></td>
<td>MTX: 20 mg/wk (oral)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karlsson et al., 2013</td>
<td></td>
<td>IFX: 3 mg/kg at wks 0, 2, 6 and every 8 wks thereafter (intravenously)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levitsky et al., 2017</td>
<td></td>
<td>G2a: Obese (BMI ≥ 30) subpopulation of G2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWEFOT</td>
<td></td>
<td>G2b: Overweight (BMI &lt;25-29.9) subpopulation of G2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country and setting</td>
<td></td>
<td>G2c: Normal (BMI &lt;25) subpopulation of G2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden; multicenter</td>
<td></td>
<td>N:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td>G1: 130</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td>G1a: 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N</td>
<td></td>
<td>G1b: 22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>258</td>
<td></td>
<td>G1c: 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of study</td>
<td></td>
<td>G2: 128</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 mos (2 yr followup)</td>
<td></td>
<td>G2a: 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2a: 26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2c: 53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age, yrs</td>
<td></td>
<td>G1: 52.9 (SD, 13.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G2: 51.1 (SD, 13.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior CS use, %</td>
<td></td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX naïve, %</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior csDMARD use, %</td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-mo Sharp score, mean</td>
<td></td>
<td>67.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF seropositive, %</td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20 response, %</td>
<td></td>
<td>ITT population:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 12 mos</td>
<td></td>
<td>G1: 28.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS27 disease activity:</td>
<td></td>
<td>G2: 42.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RR 1.48 (95% CI 1.06 to 2.08; p=0.0266)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Author, year, study name, if applicable</td>
<td>van Vollenhoven et al., 2009; Eriksson et al., 2013; van Vollenhoven et al., 2012; Rezaei et al., 2013; Eriksson et al., 2016; Levitsky et al., 2015; Karlsson et al., 2013; SWEFOT (continued)</td>
<td>Sex, % female: 76.7</td>
<td>Erosive disease, %: NR</td>
<td>mITT population:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % white: NR</td>
<td></td>
<td>G1: 45.4; G2: 59.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Race, % black: NR</td>
<td></td>
<td>RR 1.31 (95% CI, 1.03 to 1.66; p=0.0257)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethnicity, % Latino: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR50 response, %:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ITT population:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR 1.71 (95% CI, 1.02 to 2.86; p=0.0424)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mITT population:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR 1.43 (95% CI, 1.06 to 1.93; p=0.0226)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACR70 response, %:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ITT population:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR 1.69 (95% CI, 0.77 to 3.73; p=0.2044)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mITT population:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR 1.83 (95% CI, 1.12 to 2.98; p=0.0156)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS28 remission, %:</td>
</tr>
</tbody>
</table>

C-174
<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, year, study name, if applicable</td>
<td>van Vollenhoven et al., 2009&lt;sup&gt;110&lt;/sup&gt;; Eriksson et al., 2013&lt;sup&gt;121&lt;/sup&gt; van Vollenhoven et al., 2012&lt;sup&gt;112&lt;/sup&gt;; Rezaei et al., 2013&lt;sup&gt;123&lt;/sup&gt; Eriksson et al., 2016&lt;sup&gt;124&lt;/sup&gt; Levitsky et al., 2015&lt;sup&gt;125&lt;/sup&gt; Karlsson et al., 2013&lt;sup&gt;126&lt;/sup&gt; SWEFOT (continued)</td>
<td>Achieved remission at least 3 months after initiation, %: G1a: 15 G1b+G1c: 32 G2a: 42 G2b+G2c: 35</td>
<td><strong>Sharp score:</strong> NR</td>
<td><strong>HAQ:</strong> NR</td>
<td><strong>SF-36:</strong> NR</td>
<td>At 9 mos Achieved remission at least 3 months after initiation %: G1a: 0 G1b+G1c: 27 G2a: 33 G2b+G2c: 41 G1a vs G2a, P = 0.021 G1a vs G1b+G1c, P = 0.017</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong> Verschueren, et al., 2015&lt;sup&gt;1&lt;/sup&gt; Verschueren, et al., 2015&lt;sup&gt;2&lt;/sup&gt; Verschueren, et al., 2017&lt;sup&gt;3&lt;/sup&gt;</td>
<td>CareRA</td>
<td>Patients with RA defined by ACR criteria with disease duration ≤1yr. and DMARDs/glucocorticoid naïve</td>
<td><strong>Interventions, dose:</strong> G1: COBRA Classic (high-risk patients) • MTX (15 mg/wk) + SSZ (2 g/d) + PRED (60 mg/d tapered to 7.5 mg/d from wk 7) G2: COBRA Slim (high-risk patients) • MTX (15 mg/wk) + PRED (30 mg tapered to 5 mg from wk 6) G3: COBRA Avant-Garde (high-risk patients) • MTX (15 mg/wk) + LEF (10 mg/d) + PRED (30 mg tapered to 5 mg from wk 6) G4: MTX tight step up (low-risk patients) • MTX (15 mg/wk), no steroids allowed G5: COBRA Slim (low-risk patients) • MTX (15 mg/wk) + PRED (30 mg tapered to 5 mg from wk 6)</td>
<td><strong>Mean disease duration, wks:</strong> 1.8-3.2</td>
<td><strong>At 52 wks</strong> DAS28 (CRP) disease activity, mean change (SD) G1: 2.5 (SD, 1.5) G2: 2.3 (SD, 1.4) G3: 2.3 (SD, 1.5) G4: 2.1 (SD, 1.7) G5: 2.1 (SD, 1.9)</td>
<td>G1 vs. G2 vs. G3 p=0.329 G4 vs. G5 p=0.990</td>
</tr>
<tr>
<td><strong>Country, Clinical Setting:</strong> Belgium, rheumatology centers (academic, hospital, and private)</td>
<td></td>
<td></td>
<td><strong>Baseline DAS28(CRP), mean:</strong> 4.5-5.0</td>
<td><strong>Good EULAR response, %</strong> G1: 67.3 G2: 68.4 G3: 67.7 G4: 57.4 G5: 60.5</td>
<td>G1 vs. G2 vs. G3 p=0.995 G4 vs. G5 p=0.771</td>
<td><strong>Overall:</strong> G1: 15.3 G2: 15.3 G3: 10.8 G4: 14.9 G5: 16.3</td>
</tr>
<tr>
<td><strong>Study Design:</strong> RCT</td>
<td></td>
<td></td>
<td><strong>Baseline HAQ, mean:</strong> 0.9-1.2</td>
<td><strong>Moderate EULAR response, %</strong> G1: 84.7 G2: 88.8 G3: 88.2 G4: 78.7 G5: 76.7</td>
<td>G1 vs. G2 vs. G3 p=0.654 G4 vs. G5 p=0.822</td>
<td><strong>Discontinuation because of AEs</strong> NR</td>
</tr>
<tr>
<td><strong>Overall N:</strong> 379</td>
<td></td>
<td></td>
<td><strong>MTX naïve, %:</strong> 100</td>
<td><strong>Discontinuation because of AEs</strong> NR</td>
<td></td>
<td><strong>Patient adherence</strong> 69.4</td>
</tr>
<tr>
<td><strong>Study Duration:</strong> 2 yrs</td>
<td></td>
<td></td>
<td><strong>MTX inadequate responders:</strong> 0</td>
<td><strong>DAS28 &lt;2.6 remission, %</strong> G1: 64.3 G2: 60.2 G3: 62.4 G4: 57.4 G5: 67.4</td>
<td></td>
<td><strong>Itch and Rash</strong> G1: 4.1 G2: 3.1 G3: 1.1 G4: 6.4 G5: 4.7</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Verschueren, et al., 2015&lt;sup&gt;94&lt;/sup&gt;</td>
<td>PRED: Dose tapered down wkly except for the lowest dose (7.5 mg in G1 and 5 mg in G2/3) which was maintained until wk 28. After that PRED was tapered. Mean PRED dose at 52 wks was 4.9 mg/d (SD, 1.6)</td>
<td>N:</td>
<td>Change in SHS from baseline, mean (SD)</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>versus</td>
<td>G1: 98</td>
<td>G2: 98</td>
<td>G3: 93</td>
<td>G4: 47</td>
<td>G5: 43</td>
<td>G1: 0.3 (SD, 0.5)</td>
</tr>
<tr>
<td>CareRA (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2 vs. G3 p=0.819</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean age, yrs:</td>
<td>51.2-53.2</td>
<td>HAQ change according to ITT analysis after LOCF imputation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sex, % female:</td>
<td>64.3-80.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Race, % white:</td>
<td>NR</td>
<td>SF-36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At 16 wks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS disease activity, change from baseline</td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Author, yr, Study Name:</strong></td>
<td>Verschueren, et al., 2015&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>Good EULAR response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verschueren, et al., 2015&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>G1: 79.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verschueren, et al., 2015&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>G2: 79.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verschueren, et al., 2017&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>G3: 76.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CareRA (continued)</td>
<td></td>
<td></td>
<td>G4: 44.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5: 58.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2 vs. G3 p=0.844</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4 vs. G5 p=0.202</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 v G2 difference (95% CI): 0.0% (-11.3% to 11.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 v G3 difference (95% CI): -3.0% (-14.7% to 8.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Moderate EULAR response, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 98.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 95.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 93.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 72.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5: 86.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2 vs. G3 p=0.320</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4 vs. G5 p=0.111</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 v G2 difference (95% CI): 2.1% (-3.6% to 8.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 v G3 difference (95% CI): -2.3% (-9.6% to 4.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DAS remission, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 70.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 73.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G3: 68.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4: 46.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G5: 65.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 vs. G2 vs. G3 p=0.713</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G4 vs. G5 p=0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1 v G2 difference (95% CI): -3.1% (-15.4% to 9.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2 v G3 difference (95% CI): -5.4% (-18.0% to 7.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, yr, Study Name: Verschueren, et al., 2015</td>
<td>Study Population Summary: CareRA (continued)</td>
<td>Interventions and Patient Characteristics: SHS</td>
<td>Baseline Disease and Treatment Characteristics: HAQ mean change from baseline</td>
<td>Health Outcomes: G1: 0.8 (SD, 0.6), G2: 0.6 (SD, 0.6), G3: 0.7 (SD, 0.6), G4: 0.40 (SD, 0.62), G5: 0.58 (SD, 0.64), G1 vs. G2 vs. G3: p=0.081, G4 vs. G5: p=0.267, G1 vs. G2 difference (95% CI): 0.2 (0.02 to 0.37), G2 vs. G3 difference (95% CI): 0.1 (~0.17 to 0.19), HAQ score of 0 (no functional impairment), %: G1: 45.9, G2: 42.9, G3: 48.9, G4: 23.4, G5: 51.2, G1 vs. G2 vs. G3: p=0.7, G4 vs. G5: p=0.006, SF-36: NR</td>
<td>Adverse Events (%): ROB Rating: NR</td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Adults (aged ≥ 18 yrs) with disease duration ≤ 2 yrs, at least 12 tender and 10 swollen joints, CRP ≥ 0.45 mg/dl, rheumatoid factor and/or anti-CCP-2 antibodies seropositivity, and radiographic evidence of bone erosions; patients were either MTX-naive at study entry or had previous exposure of ≤10 mg/wk for ≤3 wks but not within 3 mos prior to consenting to participate</td>
<td>Interventions, dose: G1: • ABA: ~10 mg/kg on days 1, 15, 29, and every 4 wks thereafter (intravenous) • MTX: 7.5 mg/wk, 15 mg/wk at wk 4, and 20 mg/wk at wk 8 thereafter G2: • Placebo • MTX: 7.5 mg/wk, 15 mg/wk at wk 4, and 20 mg/wk at wk 8 thereafter</td>
<td>Mean disease duration, mos: 6.2-6.7 DAS28 (CRP), mean: 6.3 HAQ-DI, mean: 1.7 Prior CS use, %: 49.0-51.2 Prior csDMARD use, %: HCO: 1.6-2.0 SSZ: 0-1</td>
<td>At 1 yr DAS28 (CRP) disease activity: G1: -3.22 (SE 0.09) G2: -2.49 (SE 0.09) p&lt;0.001 ACR20 response, %: NR ACR50 response, %: G1: 57.4 G2: 42.3 p&lt;0.001 ACR70 response, %: G1: 42.6 G2: 27.3 p&lt;0.001 DAS28 (CRP) remission (&lt;2.6), %: G1: 41.4 G2: 23.3 p&lt;0.001</td>
<td>Overall AEs: G1: 84.8 G2: 83.4 SAEs: G1: 7.8 G2: 7.9 Overall discontinuation: G1: 9.4 G2: 10.3 Discontinuation due to AEs: G1: 3.1 G2: 4.3 Discontinuation due to SAEs: G1: 1.2 G2: 1.2 Discontinuation due to lack of efficacy: G1: 0.0 G2: 3.2</td>
<td>Low (ACR response, DAS28 remission, LDA, radiographic outcome, discontinuation, AEs); Medium (HAQ-DI, SF-36)</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Multinational,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country, Clinical Setting:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical Setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall N</td>
<td>509</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of study</td>
<td>1 yr (1-yr open label extension)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name:</td>
<td>Westhovens et al., 2009(^1); Wells et al., 2011(^2); Bathon et al., 2011(^3); Smolen et al., 2015(^4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AGREE (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex, % female:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 76.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 78.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % white:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 78.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 86.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, % black:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity, % Latino:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6-2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0-0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX naïve, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX inadequate responders, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biologic non-responders, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF seropositive, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>96.1-96.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-CCP-2 positive, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85.8-92.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Genant-modified Sharp score, mean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiographic evidence of bone erosions %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genant-modified Sharp score change in total score, mean:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 1.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No radiographic progression (total ≤ 0):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 61.2% (95% CI, 55.0-67.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 52.9% (95% CI, 46.6-59.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference: 8.3% (95% CI, -1.0 to 17.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achieved change of ≥0.3 units, %:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 71.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 62.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted mean change from baseline:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: -0.96 (SE 0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: -0.76 (SE 0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental component, mean change from baseline:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 8.15 (SE 0.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 6.34 (SE 0.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical component, mean change from baseline:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 11.68 (SE 0.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 9.18 (SE 0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient adherence:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific AEs, n:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 1 (pancreatic)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper respiratory infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious infections (not including pneumonia)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellulitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudomonal lung infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Characteristics</td>
<td>Study Population Summary</td>
<td>Interventions and Patient Characteristics</td>
<td>Baseline Disease and Treatment Characteristics</td>
<td>Health Outcomes</td>
<td>Adverse Events (%)</td>
<td>ROB Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Author, yr, Study Name: Westhovens et al., 2009; Wells et al., 2011; Batson et al., 2011; Smolen et al., 2015</td>
<td></td>
<td></td>
<td></td>
<td>At 6 mos DAS28 (CRP) remission (&lt; 2.6), %:</td>
<td>Post operative lung infection</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 31.4</td>
<td>G1: 1</td>
<td>G2: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 17.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Genant-modified Sharp score Change in total score, mean:</td>
<td>Breast cellulitis/ staphylococcal infection</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.47</td>
<td>G1: NR</td>
<td>G2: 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in erosion score, mean:</td>
<td>Other infections (not specified)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.40</td>
<td>G1: 132</td>
<td>G2: 139</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in joint-space narrowing, mean:</td>
<td>Infusion/injection site reactions, n</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 0.08</td>
<td>G1: 16</td>
<td>G2: 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: 0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dizziness, n</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 5</td>
<td>G2: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Most frequently reported adverse events</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Nausea: &gt; 10% pts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Upper respiratory tract infection: &gt; 10% pts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Headache: &gt; 10% pts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: NR</td>
</tr>
</tbody>
</table>
## Study Characteristics

<table>
<thead>
<tr>
<th>Author, yr, Study Name:</th>
<th>Study Population Summary</th>
<th>Interventions and Patient Characteristics</th>
<th>Baseline Disease and Treatment Characteristics</th>
<th>Health Outcomes</th>
<th>Adverse Events (%)</th>
<th>ROB Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westhovens et al., 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pregnancy (protocol violations), n</td>
<td></td>
</tr>
<tr>
<td>Wells et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1: 2</td>
<td>G2: NR</td>
</tr>
<tr>
<td>Bathon et al., 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spontaneous abortion between days 1 and 30 after 1 infusion of ABA, n</td>
<td>G1: 1</td>
</tr>
<tr>
<td>Smolen et al., 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G2: NR</td>
<td></td>
</tr>
</tbody>
</table>

a Five of our included studies reported MRI progression as an outcome evaluating high-dose corticosteroids, csDMARDs, TNF biologics, non-TNF biologics, and combinations and therapy strategies.

b The C-EARLY study’s randomized sample was 879, but baseline characteristics reflect the full analysis set of 868 patients, except for the proportion of systemic CS users, which was based on the safety set of 876 patients, and radiographic data, which used the radiographic set of 691 patients.

c Efficacy outcomes in the C-EARLY study were analyzed for the full analysis set of 868 patients, except for radiographic data which used the radiographic set of 691 patients.

d AE outcomes in the C-EARLY study were analyzed for the safety set of 876 patients, with the exceptions of overall discontinuation and discontinuation due to AEs, which were based on the randomized sample of 879 patients.

e Arm-specific data for the C-EARLY study’s specific AEs presented in this appendix (e.g., nausea) were only available on ClinicalTrials.gov.

f Of the two deaths occurring in the C-EARLY study’s CZP + MTX arm, one was caused by a stroke not considered related to study medication, and the other was a case of disseminated, non-characterized, mycobacterium infection primarily located in the peritoneum with acute respiratory distress, considered to be study medication related. The one death occurring in the MTX arm (respiratory failure) was not considered related to study medication.

AAT = alanine aminotransferase; ABA = abatacept; ACR = American College of Rheumatology (20/50/70 = 20%/50%/70% improvement); ADA = adalimumab; AE = adverse event (S = serious); ALT = alanine transaminase; ANCOVA = analysis of covariance; aOR = adjusted odds ratio; AP = alkaline phosphatase; AST = aspartate aminotransferase; BRAF-MDQ = Bristol Rheumatoid Arthritis Fatigue – Multidimensional Questionnaire; CCP = cyclic citrullinated peptide; CI = confidence interval; CRP = C-reactive protein; CS = corticosteroid; csDMARD = conventional synthetic DMARD; CZP = certolizumab pegol; DAS = Disease Activity Score (based on 44 joints); DAS28 = Disease Activity Score based on 28 joints; DMARD = disease-modifying antirheumatic drug (cs = conventional synthetic); ESR = erythrocyte sedimentation rate; ETN = etanercept; EQ-5D = EuroQoL standardized instrument; EULAR = European League against Rheumatism; Fig. = figure; G = group; GOL = golimumab; HAQ = Health Assessment Questionnaire (DI =...
### Appendix D.

**Risk of Bias Ratings and Rationales for Included Studies**

#### Appendix Table D-1. Risk of bias ratings for randomized controlled trials

<table>
<thead>
<tr>
<th>Study</th>
<th>ROB Rating(s)</th>
<th>Rationale for Rating(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGREE, 2009-15, 128-131</td>
<td>Low (ACR response, DAS28 remission, LDAS, radiographic outcomes, AEs)</td>
<td>A Low rating applies to ACR response, DAS28 remission, LDAS, radiographic outcomes, and AEs. To handle missing data, NRI was used for ACR response, DAS28 remission, and LDAS; multiple imputation was used for radiographic outcomes; and modified ITT was used for harms, such that all patients receiving one or more ABA dose were analyzed. A Medium rating applies to HAQ-DI and SF-36 outcomes because they were measured using as-observed data, but missing data were minimal for both.</td>
</tr>
<tr>
<td>ASPIRE, 2004-9, 106, 107, 157</td>
<td>Medium (HAQ-DI, SF-36)</td>
<td>ITT analysis probably not used; only patients with data after week 30 were included. However, overall attrition was fairly low at 15%.</td>
</tr>
<tr>
<td>AVERT, 2015</td>
<td>Medium</td>
<td>Attrition not described, and unable to tell if ITT was used</td>
</tr>
<tr>
<td>BARFOT Study #1, 2003</td>
<td>High</td>
<td>Treatment contamination across groups; PNL arm could have received PNL alone or PNL + MTX, and SSZ arm could have received SSZ alone or SSZ + PNL. No reporting of how findings may have differed following monotherapy vs. combination treatment within treatment arms. High overall and differential attrition also raise concern about ROB. Also, large baseline between-group differences in RF-positivity and Larsen score, such that T1 (the PNL arm) was significantly more likely than T2 (the SSZ arm) to be RF-positive and have greater radiographic damage at baseline. Statistical analyses did nothing to adjust for these differences or determine whether they could have affected the study findings.</td>
</tr>
<tr>
<td>BARFOT Study #2, 2005-14</td>
<td>Medium (1, 2, and 10-year outcomes [KQs 1-3])</td>
<td>A Medium rating applies to 1, 2, and 10-year outcomes (KQs 1-3). Open-label design introduced ROB because patients could have switched treatments based on knowledge of randomized assignments. Only radiographic outcomes measured blindly. Choice of DMARDs prescribed was similar between PNL and no-PNL arms, despite being left up to treating physicians. The significant between-group differences in NSAID and intra-articular injection use over the study's first 2 years probably not a ROB concern, but more likely reflect differences in treatment effectiveness. LOCF ITT analysis used for efficacy outcomes, except radiographic outcomes, for which completers analysis was used because investigators deemed amount of missing data minimal. No-PNL group was significantly older than the PNL group, but statistical analysis adjusted for age as a covariate. A High rating applies to 4-year outcomes (KQs 1-3) because of potential bias from high overall attrition (40%) resulting from investigator exclusion of patients and self-selection of patients into 2-year continuation study, plus attrition between 2-4 years. Baseline characteristics of the retained 4-year sample appear similar to the original study sample’s, but risk of attrition bias is still high.</td>
</tr>
<tr>
<td>Study</td>
<td>ROB Rating(s)</td>
<td>Rationale for Rating(s)</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>BeSt, 2005-16</td>
<td>Low (1-5 year outcomes)</td>
<td>Open-label design with blinded assessment for all outcomes. ITT method not specified except for DAS at 4-year timepoint and all 10-year outcomes (multiple imputation and GEE). Protocol deviation of 70 patients (14% overall) as a potential source of ROB seems unlikely because between-group differences in deviation were not significant (p=0.11), and these patients were still included in ITT analysis. Low overall and differential attrition at 1-5 year timepoints, but high enough to introduce attrition bias at 10-year timepoint (overall: 38%; differential: 3.3% to 16.5%). Therefore, a Low rating applies to outcomes measured at 1 to 5 years, while a Medium rating applies to all outcomes at the 10-year timepoint.</td>
</tr>
<tr>
<td>C-EARLY, 2017</td>
<td>Medium</td>
<td>High overall attrition for all outcomes, but especially high for work productivity outcomes that apply to KQ 2 and only reported on CT.gov (work days missed, work days with reduced productivity, interference with work productivity) due to limited availability of baseline data. Therefore, a High ROB rating applies only to KQ 2 work productivity outcomes. LOCF ITT and NRI can account for this. Potential selective outcome reporting bias affecting KQ 2-eligible PROs (e.g., fatigue, work productivity, household productivity), which were not mentioned at all in published article and only reported on CT.gov.</td>
</tr>
<tr>
<td>CAMERA-II, 2012</td>
<td>Medium</td>
<td>28% attrition is fairly high, but study not fatally flawed</td>
</tr>
<tr>
<td>CARDERA, 2008</td>
<td>Medium</td>
<td>NR whether or not care providers were masked</td>
</tr>
<tr>
<td>CareRA, 2015-17, 98, 99</td>
<td>Medium</td>
<td>No masking</td>
</tr>
<tr>
<td>COBRA, 1997-2010</td>
<td>Medium (56 week, 5 year, and most 11 year outcomes)</td>
<td>A Medium rating applies to all relevant outcomes at 56 week, 5-year, and most 11-year timepoints. High differential attrition. A High rating applies to the following 11-year outcomes: mTSS and other radiographic measures (because data only available for 112 out of 155 total patients).</td>
</tr>
<tr>
<td></td>
<td>High (11 year radiographic outcomes)</td>
<td></td>
</tr>
<tr>
<td>COBRA-light, 2014-5</td>
<td>Medium</td>
<td>24% protocol violations in COBRA and 7% in COBRA light</td>
</tr>
<tr>
<td>COMET, 2008-14</td>
<td>Medium</td>
<td>Moderate level of overall attrition. Missing outcome data was handled with LOCF for clinical outcomes and HAQ, and linear extrapolation for radiographic outcomes.</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>ITT not stated, high overall and differential attrition</td>
</tr>
<tr>
<td>Conaghan et al., 2016</td>
<td>Medium</td>
<td>High ROB rating applies to 52 weeks and 2 years. At 24 weeks, rating would be Medium because attrition is much lower. Only outcomes at 24 weeks make sense; afterwards people could switch to rescue medication and drop out rates were very high.</td>
</tr>
<tr>
<td>C-OPERA, 2016-7</td>
<td>Medium (24 week outcomes);</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High (52 week and 2 year outcomes, except discontinuation)</td>
<td></td>
</tr>
<tr>
<td>Dougados et al., 1999-2003</td>
<td>Medium</td>
<td>4 patients removed before randomization, but too small a number to affect outcome</td>
</tr>
<tr>
<td>Study</td>
<td>ROB Rating(s)</td>
<td>Rationale for Rating(s)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Durez et al., 2007&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Medium</td>
<td>Small study (N=44) with no more than 15 patients in any one arm, which could pose problems in terms of statistical power. Baseline clinical characteristics differed significantly between groups in terms of RF and anti-CCP positivity, but this did not affect findings in the sensitivity analyses conducted by authors and may have resulted simply because of small sample size. Potential selective outcome reporting bias affecting KQ 2-eligible PRO (i.e., VAS-measured pain), which was not reported in the article or on CT.gov.</td>
</tr>
<tr>
<td>Enbrel ERA, 2000-6&lt;sup&gt;14,17c&lt;/sup&gt;</td>
<td>Medium</td>
<td>High overall attrition at 2-year timepoint, and moderate overall attrition at 1-year timepoint. Also moderate differential attrition at the 2-year timepoint. Blinded outcome assessment for radiographic outcomes, but unclear if this was the case for all other eligible outcomes. Also, details about randomization were NR.</td>
</tr>
<tr>
<td>FIN-RACo, 1999-2013&lt;sup&gt;22,101,102,142-145&lt;/sup&gt;</td>
<td>Medium</td>
<td>Open label study. Minimal attrition. ITT used.</td>
</tr>
<tr>
<td>FUNCTION, 2016&lt;sup&gt;732,134&lt;/sup&gt;</td>
<td>Medium (1 year outcomes); High (2 year outcomes)</td>
<td>High overall attrition at 1 year, and much higher attrition at 2 years (47%) when taking into account the patients who were switched to rescue therapy. High ROB rating for all outcomes’ 2-year data because of attrition bias.</td>
</tr>
<tr>
<td>GUEPARD, 2009&lt;sup&gt;32&lt;/sup&gt;</td>
<td>Medium (12 week outcomes); High (52 week outcomes)</td>
<td>Open-label RCT in which only radiographic outcomes were assessed by a blind rater. Some overall attrition, but LOCF ITT analyses used to account for missing data. A Medium ROB rating applies to 12-week outcomes, but a High ROB rating for all outcomes at 52-week timepoint due to risk of contamination bias. Treatment adjustments were a potential source of contamination bias for both arms at the 52 week timepoint, since patients could be switched to different dosing and treatment regimens when low disease activity was achieved at 12 weeks and beyond (e.g., ADA+MTX --&gt; MTX alone) or in cases of insufficient response (e.g., ADA+MTX 40 mg every other week --&gt; ADA+MTX 40 mg/week --&gt; ETN). Total use of ETN in average doses was similar between arms, but between-group differences between 12-52 weeks were likely artificially lower as a result.</td>
</tr>
<tr>
<td>Haagsma et al., 1997&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Medium</td>
<td>Unclear randomization description, unclear allocation concealment</td>
</tr>
<tr>
<td>Study</td>
<td>ROB Rating(s)</td>
<td>Rationale for Rating(s)</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>HIT HARD, 2013</td>
<td>Medium (DAS28, ACR response, HAQ-DI, SF-36); High (mTSS, SHS erosion)</td>
<td>A Medium rating applies to DAS28, ACR response, HAQ-DI, and SF-36 outcomes. Factors contributing to increased ROB include overall and differential attrition at 52 weeks (with lower attrition rates at 24 weeks) and a statistically significant baseline difference between groups in age. There were also baseline differences in SF-36 physical score and SHS JSN score. A High ROB rating applies to mTSS and SHS erosion score outcomes because radiographic data were only available for 59% of ADA + MTX patients and for 55% of MTX-only patients. In fact, investigators found evidence that patients with missing radiographs differed significantly from those with complete data (for example, higher DAS28 disease activity in those with missing radiographs). Blinded outcome assessment for radiographic outcomes, but this does not attenuate ROB.</td>
</tr>
<tr>
<td>HOPEFUL 1, 2014</td>
<td>Medium</td>
<td>Some overall attrition during 26 weeks of double-blind phase, but no evidence that group similarity was unbalanced as a result. Study dosage of MTX was much lower than current approved U.S. FDA dose because this is a Japanese study done 7-8 years ago, but it seems unlikely this would have affected the magnitude of effect observed in the findings. DAS28-CRP score difference was analyzed as post-hoc outcome, but the direction and magnitude of effect seem to match those of the pre-specified DAS28(ESR) score difference. ITT methods were NRI for binary outcomes of interest (ACR20/50/70 response, DAS28 remission, % radiographic progression, HAQ-DI response, and AEs) and modified LOCF ITT for continuous outcomes (DAS28(ESR) and DAS28(CRP) scores, mTSS scores, and HAQ-DI scores).</td>
</tr>
<tr>
<td>IDEA, 2014</td>
<td>Medium</td>
<td>Unclear if allocation concealment was used</td>
</tr>
<tr>
<td>IMAGE, 2011-2012</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>IMPROVED, 2013-6</td>
<td>High</td>
<td>Only the trained research nurses conducting the DAS assessment were blinded for treatment allocation; they were not blinded for other outcome assessment. High attrition rate, ITT analysis is stated, but it's not mentioned how missing data were handled.</td>
</tr>
<tr>
<td>Marcora et al., 2006</td>
<td>Medium</td>
<td>Open-label RCT using a completers analysis with a very small sample (N=24). Still, small attrition rate (n=2 patients, or 7.7%). Unclear if outcome assessment was blinded for DAS28 change from baseline. Also unclear if arms similar in terms of erosive disease or Sharp scores.</td>
</tr>
<tr>
<td>Montecucco et al., 2012</td>
<td>Medium</td>
<td>Open label, authors report using both ITT and per-protocol analyses</td>
</tr>
<tr>
<td>NEO-RACo, 2013-5</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>OPERA, 2013-7</td>
<td>Medium</td>
<td>Low attrition rates. Study design details were well-reported and indicate a well-designed RCT. However, increased ROB from Type 2 error (i.e., potential for finding of a between-group difference when there really is none) because study was underpowered for DAS28-CRP disease response and, therefore, for all other outcomes. Treatment blinding was terminated after 1 year, and patients had their treatments reassessed based on clinician judgment through year 2. Still, similar proportions of patients were switched to triple synthetic DMARD therapy or received intra-articular injections in addition to randomized treatments.</td>
</tr>
<tr>
<td>OPTIMA, 2013-6</td>
<td>Low</td>
<td>Non-blinding of participants, outcome assessors, care providers, no ITT analysis performed</td>
</tr>
<tr>
<td>ORBIT, 2016</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>ROB Rating(s)</td>
<td>Rationale for Rating(s)</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PREMIER, 2006-1515, 103, 115-119, 149</td>
<td>Medium</td>
<td>High overall attrition. Also moderate differential attrition, but that was attributable mainly to difference in attrition because of lack of efficacy. ITT was used to account for missing data, although the specific type of ITT is not described. Blinded outcome assessment used for radiographic outcomes, but unclear if this was the case for other outcomes.</td>
</tr>
<tr>
<td>PROWD, 200816, 152</td>
<td>Medium (16 week outcomes); High (56 week outcomes, except discontinuation)</td>
<td>A High rating applies to 56 week outcomes, except study withdrawal, because of very high overall attrition and moderate differential attrition, but a Medium rating applies to 16 week outcomes, including withdrawal. Missing data were handled using LOCF for continuous outcomes, and NRI for job loss/imminent job loss.</td>
</tr>
<tr>
<td>Quinn et al., 200511</td>
<td>Medium</td>
<td>Type 2 error affected radiographic outcomes and possibly disease activity and QOL outcomes because study only statistically powered for MRI bone erosions and because of small sample size (only 10 in each arm). Method of handling dropouts not described.</td>
</tr>
<tr>
<td>SWEFOT, 2009-1710, 121-126, 168</td>
<td>Medium</td>
<td>Open-label design of this RCT creates an increased ROB in that patients more likely to discontinue conventional treatments in favor of biologics. In fact, discontinuation in conventional arm was significantly greater than in the biologic arm, &quot;accounted for mostly by participants who discontinued prematurely because of lack of effectiveness&quot;. Overall attrition exceeded 20% at 1 year timepoint, but the use of conservative NRI analysis and also modified ITT (for comparison) accounted for missing data and treatment switches. Larger overall attrition increases ROB to a borderline High level at 2 year timepoint, but statistical analyses help manage any elevated ROB.</td>
</tr>
<tr>
<td>TEAR, 2012-1320, 159</td>
<td>High</td>
<td>High attrition and modified ITT analysis not sufficient to account for attrition bias</td>
</tr>
<tr>
<td>Todoerti et al., 201010</td>
<td>Medium</td>
<td>Main flaw of this study is its open-label design, which could have introduced information bias that differentially affected how outcomes measured between groups. Randomization method unclear. Otherwise, no notable methodological issues or potential sources of bias.</td>
</tr>
<tr>
<td>tREACH, 2013-164, 146-148</td>
<td>Medium</td>
<td>Single-blinded</td>
</tr>
<tr>
<td>U-Act-Early, 2016-733, 135</td>
<td>Medium</td>
<td>High overall attrition, but ITT analyses applied to account for resulting bias. Results of ITT vs. per-protocol analyses were similar for study's primary outcome: sustained remission. Unclear how well powered the study was to detect differences in outcomes besides sustained remission (study's primary outcome). ITT methods included NRI and multiple imputation. Treatment arms were mostly similar at baseline, although male vs. female distribution differed by as much as 15% between groups. Note that 52-week data for remission only reported on study's CT.gov page</td>
</tr>
</tbody>
</table>
### Appendix Table D2. ROB ratings for observational studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>ROB Rating(s)</th>
<th>Rationale for Rating(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bili et al., 2014(^{11})</td>
<td>Retrospective cohort study</td>
<td>High</td>
<td>Not possible to draw valid conclusions from study findings because of how medication use classified. Medication use evaluated as &quot;exposure periods&quot;, and individual patients could contribute data to multiple exposure periods for different drugs. Furthermore, MTX group included MTX monotherapy and combination therapies.</td>
</tr>
<tr>
<td>ERAN, 2013(^{137})</td>
<td>Prospective cohort study</td>
<td>High</td>
<td>High risk of bias from classification of interventions. Comparisons of treatment use vs. no use provides insufficient information to draw clear usable conclusions because no-use patients would have taken at least one of seven alternative treatments (Table 1). No information on which alternative treatments patients switched to after discontinuing initial DMARD treatment.</td>
</tr>
<tr>
<td>Nijmegen RA Inception Cohort, 2009(^{26})</td>
<td>Prospective cohort study</td>
<td>High</td>
<td>High risk of selection bias for treatment discontinuation. High risk of attrition bias at 6 months (overall: 24.3%) and 12 months (overall: 41.3%; differential: 16.1%). High risk of confounding from indication.</td>
</tr>
<tr>
<td>NOR-DMARD analysis, 2012(^{28})</td>
<td>Retrospective cohort study</td>
<td>High</td>
<td>High ROB from confounding by indication, from time-varying reduction in patients being prescribed SSZ in favor of MTX, and from unbalanced use of concomitant PNL (use in MTX arm exceeded use in MTX arm).</td>
</tr>
</tbody>
</table>

DMARD = disease-modifying antirheumatic drug; MTX = methotrexate; PNL = prednisolone; RA = rheumatoid arthritis; ROB = risk of bias; SSZ = sulphasalazine; TNF = tumor necrosis factor; TNFi = TNF inhibitor(s)
## Appendix E. Strength of Evidence for Key Questions 1-4 Outcomes

### Appendix Table E-1. Disease activity, remission, radiographic outcomes, functional status, and harms (KQs 1-3)^a

<table>
<thead>
<tr>
<th>Intervention and Comparisons</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Number of Studies; # of Subjects</th>
<th>Study Limitations</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Other Limitations</th>
<th>Results</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corticosteroid vs. csDMARDs</td>
<td>Disease activity</td>
<td>Trials</td>
<td>5 RCTs: 2 double-blinded, 3 open label; N=1307</td>
<td>High: open label design and high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed for disease</td>
<td>Insufficient activity</td>
</tr>
<tr>
<td>Remission</td>
<td>Trials</td>
<td>5 RCTs: 2 double-blinded, 3 open label; N=1395</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher remission in corticosteroid + MTX vs. MTX</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trials</td>
<td>4 RCTs: 2 double-blinded, 2 open label; N=1344</td>
<td>Medium: open label design and high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for radiographic changes</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trials</td>
<td>4 RCTs: 2 double-blinded, 2 open label; N=1344</td>
<td>High: open label design and high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed for functional capacity</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>4 RCTs: 2 double-blinded, 2 open label; N=1185</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trials</td>
<td>3 RCTs: 2 double-blinded, 1 open label; N=1085</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>High dose corticosteroid vs. IFX</td>
<td>Response</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=156</td>
<td>Medium: open label design</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in ACR response</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Remission</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=156</td>
<td>Medium: open label design</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in remission</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=112</td>
<td>Medium: open label design</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in SHS scores</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Direct-ness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=156</td>
<td>Medium: open label design</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No differences in d/c due to AEs</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=156</td>
<td>Medium: open label design</td>
<td>Inconsistent</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>N/A</td>
<td>Mixed results for functional capacity</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=156</td>
<td>Medium: open label design</td>
<td>Inconsistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>Higher SAE in IFX + MTX vs. Methyl-PNL + MTX</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>High dose corticosteroid vs. csDMARD monotherapy</td>
<td>Response</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=44</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences in ACR response</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=44</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences in remission</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=44</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>Greater improvement in functional capacity in IV methyl-PNL + MTX vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=44</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences in d/c due to AEs</td>
<td>Insufficient</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Serious AEs                 | Trial       | 1 double-blinded RCT; N=44 | Medium                          | Unknown           | Direct      | Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size | None | No significant differences in SAEs | Insufficient         |</p>
<table>
<thead>
<tr>
<th>Intervention and Comparisons</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Number of Studies; # of Subjects</th>
<th>Study Limitations</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Other Limitations</th>
<th>Results</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>csDMARD monotherapy vs. csDMARD monotherapy</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=245</td>
<td>High: high attrition, and large baseline differences between groups</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences in disease activity in PNL + SSZ vs. PNL + MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Disease activity</td>
<td>Cohort</td>
<td>1 Cohort; N=1102</td>
<td>High: confounding by indication</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant difference in disease activity in SSZ vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=245</td>
<td>High: high attrition and direction of effect varies</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in remission PNL + SSZ vs. PNL + MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=245</td>
<td>High: high attrition and large baseline differences between groups</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in Larsen score in PNL + SSZ vs. PNL + MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=245</td>
<td>High: high attrition and large baseline differences between groups</td>
<td>Unknown</td>
<td>Direct</td>
<td></td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>No significant differences in functional capacity in PNL + SSZ vs. PNL + MTX</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Cohort</td>
<td>1 Cohort; N=1102</td>
<td>High: confounding by indication</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant difference in functional capacity in SSZ vs. MTX</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=245</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>Higher d/c in SSZ + PNL vs. MTX + PNL</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Cohort</td>
<td>1 Cohort; N=1102</td>
<td>High: confounding by indication</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>Higher d/c with SSZ vs. MTX</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>csDMARD combination therapy vs. csDMARD monotherapy</td>
<td>Disease activity</td>
<td>Trials</td>
<td>5 double-blinded RCTs; N=1183</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant differences in disease activity (DAS, ACR response) for comparisons of MTX + SSZ vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cohort</td>
<td>1 Cohort; N=230</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>No significant difference in disease activity for MTX + SSZ vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trials</td>
<td>5 double-blinded RCTs; N=1242</td>
<td>Medium: high attrition</td>
<td>Inconsistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and optimal information size not met</td>
<td>None</td>
<td>Mixed results for radiographic changes</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trials</td>
<td>6 double-blinded RCTs; N=1347</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>N/A</td>
<td>No significant differences in functional capacity for comparisons of MTX + SSZ vs. MTX at 1 year or 5 years. No difference in functional capacity for comparisons of PNL + MTX + SSZ + HCQ vs. MTX or SSZ</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>6 double-blinded RCTs; N=1347</td>
<td>Medium: open label design and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Cohort</td>
<td>1 Cohort; N=230</td>
<td>High: high attrition and high risk of selection bias for treatment discontinuation and confounding by indication</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>-------------------</td>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>csDMARD combination therapy vs. csDMARD monotherapy (continued)</td>
<td>Serious AEs</td>
<td>Trials</td>
<td>6 double-blinded RCTs; N =1347</td>
<td>Medium: open label design, and high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td>csDMARD plus TNF biologic vs. TNF biologic</td>
<td>Response</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>N/A</td>
<td>Higher ACR50 response for comparison of ADA + MTX vs. ADA</td>
<td>Moderate</td>
</tr>
<tr>
<td>ADA or ADA vs. MTX</td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Higher remission for ADA + MTX vs. ADA</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Lower modified Sharp Score change for ADA + MTX vs. ADA</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>N/A</td>
<td>Greater improvement in functional capacity in ADA + MTX vs. ADA</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=799</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>csDMARD plus Non-TNF biologic vs. Non-TNF biologic</td>
<td>Response</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=351</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=351</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=351</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=351</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=351</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Direct-ness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>TCZ + MTX vs. TCZ</td>
<td>Disease activity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for disease activity (DAS) for TCZ + MTX vs. TCZ</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Higher remission for TCZ + MTX vs. TCZ</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Lower Sharp score changes in TCZ + MTX vs. TCZ</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for functional capacity at 52 weeks for TCZ + MTX vs. TCZ</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>---------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>csDMARDs vs. tsDMARDs</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher DAS and ACR50 response for TOF + MTX vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher remission for TOF + MTX vs. TOF or MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>Lower Sharp score changes with TOF compared with TOF + MTX or MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>csDMARDs vs. tsDMARDs</td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>N/A</td>
<td>No difference in functional capacity between TOF + MTX vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=108</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, and not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>TNF biologic plus csDMARD vs. csDMARD</td>
<td>ADA + MTX vs. MTX</td>
<td>Disease activity</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label; N=2485</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Higher ACR50 response with ADA + MTX vs. MTX</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label; N=2485</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Higher remission with ADA + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label; N=2485</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Lower Sharp score changes for ADA + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label N=2485</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Greater improvement in functional capacity for ADA + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label; N=2485</td>
<td>Medium: high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trials</td>
<td>5 RCTs: 3 double-blind, 2 open label; N=2485</td>
<td>Medium: high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>---------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>CZP + MTX vs. MTX</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double blinded RCT; N=879</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>Higher ACR50 response at 52 wks for CZP + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Remission Trials</td>
<td></td>
<td>2 double-blinded RCT; N=1195</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>Higher DAS remission for CZP + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes Trials</td>
<td></td>
<td>2 double-blinded RCT; N=1195</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>Lower mTSS change for CZP + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Functional capacity Trials</td>
<td></td>
<td>2 double-blinded RCT; N=1195</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>Greater improvement in HAQ-DI in CZP + MTX vs. MTX group at 52 weeks</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs Trials</td>
<td></td>
<td>2 double-blinded RCT; N=1195</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Serious AEs Trials</td>
<td></td>
<td>2 double-blinded RCT; N=1195</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: optimal information size not met, and large CIs</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td><strong>ETN + MTX vs. MTX and ETN vs. MTX</strong></td>
<td>Disease activity</td>
<td>Trials</td>
<td>3 double-blinded RCTs; N=2000</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Higher ACR20 response for ETN + MTX and ETN vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td>Remission</td>
<td>Trial</td>
<td></td>
<td>1 double-blinded RCT; N=542</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher remission for ETN + MTX and ETN vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trials</td>
<td></td>
<td>2 double-blinded RCTs; N=1174</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Lower Sharp score change for ETN + MTX and ETN vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trials</td>
<td></td>
<td>3 double-blinded RCTs; N=2000</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs</td>
<td>None</td>
<td>Mixed results for functional capacity</td>
<td>Low</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trials</td>
<td></td>
<td>3 double-blinded RCTs; N=2000</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trials</td>
<td></td>
<td>2 double-blinded RCTs; N=2000</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>---------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>IFX + MTX vs. MTX</td>
<td>Disease activity</td>
<td>Trials</td>
<td>3 double-blinded RCTs; N=1113</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>Mixed results for ACR response</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trials</td>
<td>3 double-blinded RCTs; N = 1113</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Higher remission for IFX + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1069</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>Mixed results for radiographic progression</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trials</td>
<td>3 double-blinded RCTs; N=1113</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Greater functional capacity with IFX + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N = 1093</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N = 1093</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>TNF biologic vs. csDMARD combination therapy</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=161</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences in DAS</td>
<td>Insufficient</td>
</tr>
<tr>
<td>ADA + MTX vs. MTX + PRED + HCQ + SSZ</td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=161</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences in remission</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=161</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No difference in radiographic score progression</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-----------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>ADA + MTX vs. MTX + PRED + HCQ + SSZ (continued)</td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=161</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No difference in functional capacity</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=161</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>IFX + MTX vs. MTX + SSZ + HCQ</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=258</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Increased ACR50 response for IFX + MTX vs. MTX + SSZ + HCQ</td>
<td>Low</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=258</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=258</td>
<td>Medium</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------------------</td>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>IFX + MTX + SSZ + HCQ + PRED vs. MTX + SSZ + HCQ + PRED</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences in ACR responses</td>
<td>Low</td>
</tr>
<tr>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences in remission</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences in radiographic score progression</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No difference in functional capacity</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>D/C due to Aes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=99</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Non-TNF biologic vs. csDMARD monotherapy</td>
<td>Disease activity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=860</td>
<td>Medium: high attrition, and large baseline differences between groups</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Improved disease activity with ABA + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N = 860</td>
<td>Medium: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Higher remission rates for ABA + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trials</td>
<td>1 double-blinded RCT; N=509</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Lower Genant-modified Sharp scores in ABA + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N = 860</td>
<td>Medium: high attrition</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>Mixed results for functional capacity between ABA + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td></td>
<td>1 double-blinded RCT; N=509</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td></td>
<td>1 double-blinded RCT; N=509</td>
<td>Medium: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No differences</td>
<td>Low</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other Limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>RIT + MTX vs. MTX</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Improved disease activity with RIT + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher remission with RIT + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Lower Genant-modified Sharp scores in RIT + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>Greater improvement in functional capacity in RIT + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Moderate</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>----------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>TCZ + MTX vs. MTX</td>
<td>Disease activity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for disease activity (DAS)</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Higher remission for TCZ + MTX vs. MTX</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Radiographic changes</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium: large baseline differences between groups</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>Lower Sharp score changes in TCZ + MTX vs. MTX</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Inconsistent: direction of effect varies</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for functional capacity at 52 weeks for TCZ + MTX vs. MTX</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Serious AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=1479</td>
<td>Medium</td>
<td>Consistent</td>
<td>Direct</td>
<td>Precise</td>
<td>None</td>
<td>No significant differences</td>
<td>Moderate</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>--------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>TNF vs. Non-TNF</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 open label RCT; N=329</td>
<td>High: no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences in DAS for RIT vs. ADA or ETN</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Remission</td>
<td>Trial</td>
<td>1 open label RCT; N=329</td>
<td>High: no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No significant differences in remission for RIT vs. ADA or ETN</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 open label RCT; N=329</td>
<td>High: no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Precise</td>
<td>N/a</td>
<td>Greater improvement in functional capacity in RIT vs. TNF biologic (ADA or ETN)</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 open label RCT; N=329</td>
<td>High: no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>No differences</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>TNF vs. Non-TNF (continued)</td>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 open label RCT; N=329</td>
<td>High: no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>Higher for RIT vs. ADA or ETN</td>
<td>Insufficient</td>
</tr>
</tbody>
</table>

**Combination strategies**

| 1-Sequential monotherapy starting with MTX vs. 2-step-up combination therapy vs. 3-combination with high-dose tapered prednisone vs. 4-combination therapy with infliximab | Disease activity | Trial | 1 double-blinded RCT; N=508 | Low | Unknown | Direct | Imprecise: large CIs cross appreciable benefits or harms | None | Lower disease activity scores for 3 (combination therapy with high dose prednisone) and 4 (combination therapy with IFX) than 1 (sequential DMARD therapy) or 2 (step-up therapy) at one year but no differences at 4 yrs and 10 years. | Moderate |

<p>| Remission | Trial | 1 double-blinded RCT; N=508 | Low | Unknown | Direct | Imprecise: large CIs cross appreciable benefits or harms | None | No differences in remission at 4 yrs and 10 years | Moderate |</p>
<table>
<thead>
<tr>
<th>Intervention and Comparisons</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Number of Studies; # of Subjects</th>
<th>Study Limitations</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Other Limitations</th>
<th>Results</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Sequential monotherapy starting with MTX vs. 2-step-up combination therapy vs. 3-combination with high-dose tapered prednisone vs. 4-combination therapy with infliximab (continued)</td>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=508</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Lower Sharp/van der Heijde radiographic changes in groups 3-combination therapy with high dose prednisone) and 4(combination therapy with IFX) than 1 (sequential DMARD therapy) or 2 (step up therapy) at 5 years but no differences at 10 years.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=508</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms, not enough events to meet optimal information size</td>
<td>None</td>
<td>Greater functional capacity in groups 3 (combination therapy with high dose prednisone) and 4 (combination therapy with IFX) than 1 (sequential DMARD therapy) or 2 (step up therapy) at 12 months, but no significant difference at 2 years, 5 years or 10 years.</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=508</td>
<td>Low</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>1-immediate MTX + ETN vs. 2-immediate MTX + SSZ + HCQ vs. 3-step up MTX to combo MTX + ETN vs. 4-step up MTX to combo MTX + SSZ + HCQ</td>
<td>Disease activity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>The 2 immediate groups (groups 1 and 2) had improved disease activity compared with step up (groups 3 and 4) at 6 months, but no differences at 2 yrs</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Remission</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant changes in remission at 2 yrs</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant changes in radiographic scores at 2 yrs</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences at 48 and 102 weeks</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=755</td>
<td>High: high attrition and no ITT analysis</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Intervention and Comparisons</td>
<td>Outcome</td>
<td>Study Design</td>
<td>Number of Studies; # of Subjects</td>
<td>Study Limitations</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Other limitations</td>
<td>Results</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>ADA + MTX adjusted based on DAS vs. MTX</td>
<td>Disease Activity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=245</td>
<td>High: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No differences in ACR response at 2 yrs</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Remission</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=245</td>
<td>High: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No differences in remission at 2 yrs</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Radiographic changes</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=245</td>
<td>High: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No differences in radiographic changes at 2 yrs</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Functional capacity</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=245</td>
<td>High: high attrition</td>
<td>Inconsistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>Mixed results for functional capacity at 1 yr</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>D/C due to AEs</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=180</td>
<td>High: high attrition</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
<tr>
<td>Serious AEs</td>
<td>Trials</td>
<td>2 double-blinded RCTs; N=245</td>
<td>High: high attrition</td>
<td>Consistent</td>
<td>Direct</td>
<td>Imprecise: large CIs cross appreciable benefits or harms</td>
<td>None</td>
<td>No significant differences</td>
<td>Insufficient</td>
<td></td>
</tr>
</tbody>
</table>

*a Consistent with network meta-analysis. For the SOE for effect estimates derived from indirect comparisons only (i.e., no head to head trials), the SOE for all estimates was low. We downgraded for indirectness and precision in all cases. The NWMA model included only studies with low or unclear risk of bias, therefore we did not downgrade for study limitations. Because of the single estimate derived from the NWMA, we also did not downgrade for inconsistency.*
ABA = abatacept; ACR = American College of Rheumatology; ACR50 = American College of Rheumatology 50% improvement; ADA = adalimumab; AEs = adverse events; CI = confidence interval; csDMARD = conventional synthetic disease-modifying antirheumatic drug; CZP = certolizumab pegol; d = day; DAS = Disease Activity Score; D/C = discontinuation; DMARD = disease-modifying antirheumatic drug(s); ETN = etanercept; HAQ = Health Assessment Questionnaire; HCQ = hydroxychloroquine; IFX = infliximab; ITT = intent-to-treat; mTSS = modified Sharp/van der Heijde score; MTX = methotrexate; N = number of patients; NA = not applicable; NWMA = network meta-analysis; obs = observational; PNL = prednisolone; PRED = prednisone; RCT = randomized controlled trial; RIT = rituximab; SAE = serious adverse events; SHS = Sharp/van der Heijde score; SOE = strength of evidence; SSZ = sulfasalazine; TCZ = tocilizumab; TNF = tumor necrosis factor; tsDMARD = targeted synthetic disease-modifying antirheumatic drug; vs. = versus; yr = year.
<table>
<thead>
<tr>
<th>Intervention and Comparisons</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Number of Studies; N of Subjects</th>
<th>Study Limitations</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Reporting Bias</th>
<th>Other Limitations</th>
<th>Results</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TNF biologic plus csDMARD</strong> vs. <strong>csDMARD:</strong></td>
<td>Disease activity/radiographic change</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=171</td>
<td>High: no test of interaction for subgroup analyses; results based on regression analyses</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: study does not meet optimal information size for subgroup analyses</td>
<td>Undetected</td>
<td>None</td>
<td>Disease activity is significantly associated with radiographic change</td>
<td>Insufficient</td>
</tr>
<tr>
<td>ADA + MTX vs. MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TNF biologic vs. csDMARD:</strong></td>
<td>Age/response</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=424</td>
<td>High: no test of interaction for subgroup analyses; results based on regression analyses</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: no test of interaction for subgroup analyses</td>
<td>Undetected</td>
<td>None</td>
<td>Lower ACR response rates for older (&gt;65 years) compared with younger patients</td>
<td>Insufficient</td>
</tr>
<tr>
<td>ETN vs. MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TNF biologic plus csDMARD</strong> vs. <strong>csDMARD:</strong></td>
<td>Age/SAE</td>
<td>Trial</td>
<td>1 double-blinded RCT; N=424</td>
<td>High: no test of interaction for subgroup analyses; results based on regression analyses</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: no test of interaction for subgroup analyses</td>
<td>Undetected</td>
<td>None</td>
<td>Higher risk of serious adverse events for older (&gt;65 years) compared with younger patients</td>
<td>Insufficient</td>
</tr>
<tr>
<td>IFX + MTX vs. MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E-32
<table>
<thead>
<tr>
<th>Intervention and Comparisons</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Number of Studies; N of Subjects</th>
<th>Study Limitations</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Reporting Bias</th>
<th>Other Limitations</th>
<th>Results</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF biologic vs. csDMARD combo therapy:</td>
<td>Obesity/remission</td>
<td>Trial</td>
<td>1 open-label RCT; N=260</td>
<td>High: no test of interaction for subgroup analyses; results based on regression analyses</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: Study does not meet optimal information size for subgroup analyses</td>
<td>Undetected</td>
<td>None</td>
<td>Obesity is significantly associated with lower rates of remission</td>
<td>Insufficient</td>
</tr>
<tr>
<td>IFX + MTX vs. csDMARD combo</td>
<td>Obesity/response</td>
<td>Trial</td>
<td>1 open-label RCT; N=260</td>
<td>High: no test of interaction for subgroup analyses; results based on regression analyses</td>
<td>Unknown</td>
<td>Direct</td>
<td>Imprecise: Study does not meet optimal information size for subgroup analyses</td>
<td>Undetected</td>
<td>None</td>
<td>Obesity is significantly associated with lower rates of response</td>
<td>Insufficient</td>
</tr>
</tbody>
</table>

ACR = American College of Rheumatology; ADA = adalimumab; csDMARD = conventional synthetic disease-modifying antirheumatic drug(s); ETN = etanercept; IFX = infliximab; IV = intravenous; MTX = methotrexate; N = number of patients; NA = not applicable; obs = observational; RCT = randomized controlled trial; SAE = serious adverse events; TNF = tumor necrosis factor; vs. = versus.
Appendix F.
Eligible Clinical and Self-Reported Scales and Instruments Commonly Used in Eligible Studies of Drug Therapy for Rheumatoid Arthritis

Introduction

This appendix provides a brief overview of the various scales and self-reported measures that investigators used to assess outcomes in all the studies reviewed in this systematic review. The main outcome categories involve radiologic assessments of joint damage (erosion or narrowing) and various instruments that patients or subjects used to report on functional capacity or quality of life; the latter fall into two groups, one related to general health measures and one related to condition- or disease-specific instruments. General measures used in rheumatoid arthritis studies are described first; then the disease-specific measures used in rheumatoid arthritis studies are described separately. The new 2010 American College of Rheumatology (ACR) criteria are presented at the end of the document (Appendix Table F2).

Radiographic Measures

Radiographic assessment of joint damage in hands (including wrists) or both hands and feet are critical to clinical trials in rheumatoid arthritis. The damage can be both joint space narrowing and erosions, and the underlying construct is sometimes referred to as radiographic progression (i.e., changes, whether positive or negative) as detected by radiography and interpretation. Several approaches exist, but the two commonly used are the Sharp Score (and variants) and the Larsen Score. These and other scoring methods have recently been reviewed by Boini and Guillemín; additional citations or sources are given in the brief descriptions below.

Sharp Score and Sharp/van der Heijde Score

The Sharp Score is a means of evaluating joint damage in joints of the hands, including both erosion and joint space narrowing. Although it has undergone modifications since its introduction, the version proposed in 1985 has become the standard approach. In this method, 17 joint areas in each hand are scored for erosions; 18 joint areas in each hand are scored for joint space narrowing. The score per single joint for erosions ranges from 0 to 5 and for joint space narrowing from 0 to 4. In both cases, a higher score is worse. Erosion scores range from 0 to 170 and joint space narrowing scores range from 0 to 144. Thus, the “total Sharp Score” is the sum of the erosion and joint space narrowing scores, or 0 to 314.

The Sharp/van der Heijde (SHS) method, introduced in 1989, overcame one drawback to the Sharp Score, namely its focus on only hands, given that feet can also be involved early in rheumatoid arthritis. Therefore, the SHS method was developed to take account of erosions and joint space narrowing in both hands and feet. As with the Sharp Score, higher scores reflect worse damage. Erosion is assessed in 16 joints in each hand and 6 joints in each foot. Each joint is scored from 0 to 5 with a maximal erosion score of 160 in the hands and 120 in the feet. Joint space narrowing and subluxation are assessed in 15 joints in the hands and 6 joints in the feet. Each joint is scored from 0 to 4 with a maximal score of 120 in the hands and 48 in the feet. The erosion and joint space narrowing scores are combined to give a total SHS score with a maximum of 448 (weighted toward hands because more joints are scored).
Numerous variants on the Sharp or SHS scores have been developed, differing subtly in terms of the numbers of joints measured and other details. Generally, all the Sharp methods are very detailed assessments and the approach, although reliable and sensitive to change, is considered time-consuming and tedious. For a speedier approach, Larsen and colleagues developed a simpler approach.

**Larsen Scale for Grading Radiographs**

The Larsen Scale is an overall measure of joint damage, originally devised in the 1970s and updated most recently in the late 1990s. It produces both a score for each joint (hands and feet) and an overall score that reflects measurement and extent of joint damage. Scores range from 0 (“normal conditions,” i.e., intact bony outlines and normal joint space) to 5 (“mutilating abnormality,” i.e., original bony outlines have been destroyed), so higher scores reflect greater damage. Scores can range from 0 to 250.

**General Health Measures**

**Health Assessment Questionnaire**

The Health Assessment Questionnaire (HAQ) is a widely used self-report measure of functional capacity; it is a dominant instrument in studies of patients with arthritis (particularly trials of drugs in patients with rheumatoid arthritis), but it is considered a generic (not disease-specific) instrument. The accepted minimally clinically important difference (MCID) for HAQ-DI in RA is a change of 0.22-0.25. Other detailed information on its variations, scoring, etc., can be found at [www.cher.brown.edu/pcoc/EHAQDESCRSCRORINGHAQ372.PDF](http://www.cher.brown.edu/pcoc/EHAQDESCRSCRORINGHAQ372.PDF) (accessed for this purpose 10/3/2017) or [www.hqlo.com/content/1/1/20](http://www.hqlo.com/content/1/1/20) (accessed for this purpose 10/3/2017) and in the seminal reports by Fries et al. and Ramey et al.

The full, five-dimension HAQ consists of four domains: disability, discomfort and pain, toxicity, and dollar costs, plus death (obtained through other sources). More commonly, “the HAQ” as used in the literature refers to the shorter version encompassing the HAQ Disability Index (HAQ-DI), the HAQ pain measure, and a global patient outcome measure. The HAQ-DI is sometimes used alone.

The HAQ-DI, with the past week as the time frame, focuses on whether the respondent “is able to…” do the activity and covers eight categories in 20 items: dressing and grooming, arising, eating, walking, hygiene, reach, grip, and common daily activities. The four responses for the HAQ-DI questions are graded as follows: without any difficulty = 0; with some difficulty = 1; with much difficulty = 2; and unable to do = 3. The highest score for any component question in a category determines the category score. The HAQ-DI also asks about the use of aids and devices to help with various usual activities. Two composite scores can be calculated, one with and one without the aids/devices element; both range from 0 to 3.

The HAQ pain domain is measured on a doubly-anchored horizontal visual analog scale (VAS) of 15 cm in length; one end is labeled “no pain” (score of 0) and the other is labeled “very severe pain” (score of 100). Patients mark a spot on the VAS, and scores are calculated as the length from “no pain” in centimeters (cm) multiplied by 0.2 to yield a value that can range between 0 and 3.

With respect to interpretation, HAQ-DI scores of 0 to 1 are generally considered to represent mild to moderate disability, 1 to 2 moderate to severe disability, and 2 to 3 severe to very severe disability.
The HAQ global health status scale measures quality of life (essentially, as how the patient is feeling) with a 15 cm doubly-anchored horizontal VAS scored from 0 (very well) to 100 (very poor).

Medical Outcomes Study Short Form 36 Health Survey

The Medical Outcomes Study Short Form 36 Health Survey (SF-36) is an internationally known generic health survey instrument. Information can be found at https://www.rand.org/health/surveys_tools/mos/36-item-short-form.html (accessed for this purpose 10/3/2017) and in a large number of articles documenting its psychometric properties. It comprises 36 items in eight independent domains tapping functioning and well-being: physical functioning, role-physical, bodily pain, and general health in one grouping (physical health) and vitality, role-emotional, social functioning, and mental health in another grouping (mental health). The SF-36 provides a separate scale score for each domain (yielding a profile of health) and two summary scores, one for physical health and one for mental health. Each scale is scored from 0 to 100 where higher scores indicate better health and well-being. The accepted MCIDs for the SF-36 physical component score in RA is 4.4, and for the SF-36 mental component score, it is 3.1.

A “version 2” of the SF-36 was introduced in the late 1990s to correct some drawbacks in formatting, wording, and other issues and to update the norm-based scoring with 1998 data. It can be fielded in two versions varying by recall period: 4-week recall (the usual approach) and 1-week recall (acute). More recently, it has been tested and used for computer adaptive testing according to item response theory principles.

EuroQol EQ-5D Quality of Life Questionnaire

A third generic quality-of-life instrument is the EuroQol EQ-5D Quality of Life Questionnaire, typically known just as the EQ-5D. More information can be found at http://www.euroqol.org/ (accessed for this purpose 10/3/2017) and in key descriptive articles, one of which is about patients with rheumatoid arthritis.

The EQ-5D covers health status in five domains (three questions each): mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. It is intended for self-response but can be used in other administration modes. Each item can take one of three response levels – no problems, some moderate problems, extreme problems – identified as level 1, 2, or 3, respectively. This yields a profile of one level for each of the five domains; this is essentially a five-digit number, and no arithmetic properties attach to these values. Users can convert health states in the five-dimensional descriptive system into a weighted health state index by applying scores from EQ-5D "value sets" elicited from general population samples to the profile pattern (e.g., 1, 2, 3, 3, 1).

The EQ-5D also has a global health VAS scale (20 cm) scored from 0 to 100.

Rheumatoid Arthritis Measures

American College of Rheumatology 20/50/70

The American College of Rheumatology (ACR) criteria are concerned with improvement in counts of tender and swollen joints and several domains of health. A principal aim of these criteria is use in studies (particularly trials) of drugs for rheumatoid arthritis. More information
can be found at
https://www.rheumatology.org/Portals/0/Files/ACR%20Preliminary%20Definition%20Of%20Improvement%20In%20Rheumatoid%20Arthritis%20Manuscript.pdf and

Today, based on work done in the mid-1990s, values for clinical trial patients are defined as improvement in both tender and swollen joint counts and in three of the following: patient’s assessment of pain; patient’s global assessment of disease activity, patient’s assessment of physical function (sometimes referred to as physical disability), the physician’s global assessment of disease activity, and acute phase reactant (C-reactive protein, or CRP). The 20, 50, or 70 designations (sometimes called the ACR Success Criteria) refer to improvements in percentage terms to 20 percent, 50 percent, or 70 percent in the relevant dimensions. A physician’s global assessment of 70 percent improvement is considered remission.

Thus, patients are said to meet ACR 20 criteria when they have at least 20 percent reductions in tender and swollen joint counts and in at least three of the domains. ACR50 and ACR70 criteria are defined in a manner similar to that for ACR 20, but with improvement of at least 50 percent and 70 percent in the individual measures, respectively. The table illustrates, in a study context, how a patient might be said to have an ACR50 response.

### EULAR Response Criteria

The European League Against Rheumatism (EULAR) response criteria classify patients as good, moderate, or non-responders based on both change in disease activity and current disease activity, using either the DAS or the DAS28 (see description above). For example, to be classified as a good responder a patient must have relevant change in DAS (≥1.2) and low current disease activity (≤2.4), while a non-responder must have ≤0.6 change in DAS and high disease activity (>3.7).

The EULAR criteria have been validated in multiple clinical trials, and confirmed in an analysis of nine clinical trials that concluded a high level of agreement and equal validity between ACR and EULAR improvement classifications. Good and moderate responders showed significantly more improvement in functional capacity and significantly less progression of joint damage than patients classified as non-responders.

### Disease Activity Score

The Disease Activity Score (DAS) is an index of disease activity first developed in the mid 1980s. The history of its development and current definitions, scoring systems, and other details can be found at [https://www.das-score.nl/das28/en/](https://www.das-score.nl/das28/en/) (accessed for this purpose 10/3/2017) and in recent articles. The DAS originally included the Ritchie Articular Index (see below), the 44 swollen joint count, the erythrocyte sedimentation rate, and a general health assessment on a
Scores on the DAS can range from 1 to 9. A cut-off level of the DAS of 1.6 is equivalent with being in remission.

More recently, an index of RA disease activity using only 28 joints – the DAS28 – has been developed, focusing on joint counts for both tenderness (TJC) and swelling (SJC). It also uses either the patient’s or a physician’s global assessment (PGA) of disease activity (on a 100 mm VAS) and the erythrocyte sedimentation rate (ESR) or C-reactive protein. The formula for calculating a DAS28 score is as follows: $= (0.56 \times TJC^{1/2}) + (0.28 \times SJC^{1/2}) + (0.7 \times \ln [ESR]) + (0.014 \times PGA \ [in \ mm])$. Numerous formulas to calculate a variety of DAS and DAS28 scores exist (see the website above), such as when a global patient assessment of health is unavailable.

The DAS28 yields a score on a scale ranging from 0 to 10. A DAS28 of 2.6 is considered to correspond to remission; a DAS28 of 3.2 is a threshold for low disease activity; and a DAS28 of more than 5.1 is considered high disease activity.

**Ritchie Articular Index**

This is a long-standing approach to doing a graded assessment of the tenderness of 26 joint regions, based on summation of joint responses after applying firm digital pressure. Four grades can be used: 0, patient reported no tenderness; +1, patient complained of pain; +2, patient complained of pain and winced; and +3, patient complained of pain, winced, and withdrew. Thus, the index ranges from 0 to 3 for individual measures and 0 to 78 overall, with higher scores being worse tenderness.

Certain joints are treated as a single unit, such as the metacarpal-phalangeal and proximal interphalangeal joints of each hand and the metatarsal-phalangeal joints of each foot. For example, the maximum score for the five metacarpal-phalangeal joints of the right hand would be 3, not 15. No weights are used for different types of joints (e.g., by size), because the issue is one of measuring changes (improvements) in tenderness; this is especially relevant for rheumatoid arthritis.
Appendix Table F-2. 2010 rheumatoid arthritis criteria

Target population (Who should be tested?)

Patients who

- have at least 1 joint with definite clinical synovitis (swelling)
  - Criteria aimed at classification of newly presenting patients; patients with erosive disease typical of RA with a history compatible with prior fulfillment of the 2010 criteria should be classified as having RA; patients with longstanding disease, including those whose disease is inactive (with or without treatment) who, based on retrospectively available data, have previously fulfilled the 2010 criteria should be classified as having RA
- with the synovitis not better explained by another disease

Differential diagnoses vary among patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted

<table>
<thead>
<tr>
<th>Classification criteria for RA</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score-based algorithm:</td>
<td></td>
</tr>
<tr>
<td>- Add score of categories:</td>
<td></td>
</tr>
<tr>
<td>Joint involvement, serology, reactants, duration</td>
<td></td>
</tr>
<tr>
<td>- Differential diagnoses vary among patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted</td>
<td></td>
</tr>
<tr>
<td>- Score of ≥6/10 needed for classification of a patient as having definite RA</td>
<td></td>
</tr>
<tr>
<td>- Although patients with a score of &lt;6/10 are not classifiable as having RA, their status can be reassessed and the criteria might be fulfilled cumulatively over time</td>
<td></td>
</tr>
</tbody>
</table>

Joint involvement

Joint involvement refers to any swollen or tender joint on examination, which may be confirmed by imaging evidence of synovitis; d Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are excluded from assessment; categories of joint distribution are classified according to the location and number of involved joints, with placement into the highest category possible based on the pattern of joint involvement

<table>
<thead>
<tr>
<th>Joint involvement</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 large joint</td>
<td>0</td>
</tr>
<tr>
<td>&quot;Large joints&quot; refers to shoulders, elbows, hips, knees, and ankles</td>
<td></td>
</tr>
<tr>
<td>2-10 large joints</td>
<td>1</td>
</tr>
<tr>
<td>1-3 small joints (with or without involvement of large joints)</td>
<td>2</td>
</tr>
<tr>
<td>&quot;Small joints&quot; refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.</td>
<td></td>
</tr>
<tr>
<td>4-10 small joints (with or without involvement of large joints)</td>
<td>3</td>
</tr>
<tr>
<td>&gt;10 joints (at least 1 small joint)</td>
<td>5</td>
</tr>
<tr>
<td>In this category, at least 1 of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (e.g., temporomandibular, acromioclavicular, sternoclavicular, etc.)</td>
<td></td>
</tr>
</tbody>
</table>
**Serology (at least 1 test result is needed for classification)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative RF and negative ACPA</td>
<td>0</td>
</tr>
<tr>
<td>Low-positive RF or low-positive ACPA</td>
<td>2</td>
</tr>
<tr>
<td>High-positive RF or high-positive ACPA</td>
<td>3</td>
</tr>
</tbody>
</table>

**Acute-phase reactants (at least 1 test result is needed for classification)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal CRP and normal ESR</td>
<td>0</td>
</tr>
<tr>
<td>Abnormal CRP or abnormal ESR</td>
<td>1</td>
</tr>
</tbody>
</table>

**Duration of symptoms**

<table>
<thead>
<tr>
<th>Duration of Symptoms</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;6 weeks</td>
<td>0</td>
</tr>
<tr>
<td>≥6 weeks</td>
<td>1</td>
</tr>
</tbody>
</table>

Appendix G. Tests of Consistency for Main Network Meta-Analyses

Main Network Meta-Analyses

We identified a total of 14 studies with a low or medium risk of bias for use in our main network meta-analyses (NWMA) comparing the efficacy of drug therapies for early rheumatoid arthritis. Those findings are presented in our main report.

Below, we present findings for our tests of consistency for specific drug comparisons.

Tests of Consistency: Models Excluding High Risk of Bias Studies

To test for consistency, we compared consistency and inconsistency models. In addition, where there were closed loops in the network diagram with both direct and indirect evidence available, we examined differences in results between direct and indirect evidence using network sidesplits. Of the closed loops in the networks, network sidesplits could not be computed for the loop consisting of Tocilizumab, Tocilizumab+MTX, and MTX because all three treatments were included in the same two trials and therefore, only direct evidence was available.

ACR50 Response

For the ACR50 outcome (see Appendix Table G-1), there was no significant difference in the consistency and inconsistency models ($\chi^2(1)=0.28$, p=0.868). Results did not differ significantly between direct and indirect evidence for Abatacept versus Abatacept plus MTX (coefficient [95% CI]=-0.09 [-0.69 to 0.52], p=0.777) or for Infliximab plus MTX versus Methylprednisolone plus MTX (coefficient [95% CI]= -0.37 [-1.99 to 1.25], p=0.653).

Remission According to Disease Activity Score

For the DAS outcome (see Appendix Table G-2), there was no significant difference in the consistency and inconsistency models ($\chi^2(1)=0.52$, p=0.772). Results did not differ significantly between direct and indirect evidence for Abatacept versus Abatacept plus MTX (coefficient [95% CI]= -0.60 [-2.31 to 1.11], p=0.491) or for Infliximab plus MTX versus Methylprednisolone plus MTX (coefficient [95% CI]= -0.44 [-2.72 to 1.84], p=0.705).

All Withdrawals

For all withdrawals (see Appendix Table G-3), there was no significant difference in the consistency and inconsistency models ($\chi^2(1)=0.43$, p=0.808). Results did not differ significantly between direct and indirect evidence for Abatacept versus Abatacept plus MTX (coefficient [95% CI]= -0.31 [-2.01 to 1.38], p=0.716) or for Infliximab plus MTX versus Methylprednisolone plus MTX (coefficient [95% CI]= 1.29 [-3.33 to 5.90], p=0.585).

Withdrawals Due to Adverse Events

For the DAS outcome (see Appendix Table G-4), there was no significant difference in the consistency and inconsistency models ($\chi^2(1)=2.86$, p=0.239). Results did not differ significantly between direct and indirect evidence for Abatacept versus Abatacept plus MTX (coefficient [95% CI]= -1.92 [-4.15 to 0.31], p=0.091) or for Infliximab plus MTX versus Methylprednisolone plus MTX (coefficient [95% CI]= 0.16 [-6.23 to 6.55], p=0.962).
### Appendix Table G-1. Table with network sidesplits: ACR50 response

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept+MTX</td>
<td>0.16</td>
<td>-0.07, 0.38</td>
<td>0.178</td>
<td>0.24</td>
<td>-0.33, 0.82</td>
<td>0.406</td>
<td>-0.09</td>
<td>-0.69, 0.52</td>
<td>0.777</td>
</tr>
<tr>
<td>Infliximab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>0.00</td>
<td>-0.51, 0.51</td>
<td>1.000</td>
<td>0.37</td>
<td>-1.17, 1.91</td>
<td>0.636</td>
<td>-0.37</td>
<td>-1.99, 1.25</td>
<td>0.653</td>
</tr>
</tbody>
</table>

ACR50 = American College of Rheumatology 50% response; CI = confidence interval; MTX = methotrexate

### Appendix Table G-2. Table with network sidesplits: Remission according to Disease Activity Score

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept+MTX</td>
<td>0.35</td>
<td>-0.26, 0.96</td>
<td>0.257</td>
<td>0.95</td>
<td>-0.66, 2.57</td>
<td>0.247</td>
<td>-0.60</td>
<td>-2.31, 1.11</td>
<td>0.491</td>
</tr>
<tr>
<td>Infliximab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>0.10</td>
<td>-0.62, 0.81</td>
<td>0.795</td>
<td>0.53</td>
<td>-1.61, 2.68</td>
<td>0.625</td>
<td>-0.44</td>
<td>-2.72, 1.84</td>
<td>0.705</td>
</tr>
</tbody>
</table>

CI = confidence interval; MTX = methotrexate

### Appendix Table G-3. Table with network sidesplits: All withdrawals

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept+MTX</td>
<td>-0.47</td>
<td>-1.08, 0.13</td>
<td>0.126</td>
<td>-0.16</td>
<td>-1.68, 1.37</td>
<td>0.839</td>
<td>-0.31</td>
<td>-2.01, 1.38</td>
<td>0.716</td>
</tr>
<tr>
<td>Infliximab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>0.00</td>
<td>-2.68, 2.68</td>
<td>1.000</td>
<td>-1.29</td>
<td>-5.04, 2.47</td>
<td>0.502</td>
<td>1.29</td>
<td>-3.33, 5.90</td>
<td>0.585</td>
</tr>
</tbody>
</table>

CI = confidence interval; MTX = methotrexate

### Appendix Table G-4. Table with network sidesplits: Withdrawals due to adverse events

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept+MTX</td>
<td>-1.48</td>
<td>-2.26, -0.70</td>
<td>&lt;0.001</td>
<td>0.44</td>
<td>-1.43, 2.32</td>
<td>0.644</td>
<td>-1.92</td>
<td>-4.15, 0.31</td>
<td>0.091</td>
</tr>
<tr>
<td>Infliximab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>-1.10</td>
<td>-4.22, 2.03</td>
<td>0.491</td>
<td>-1.26</td>
<td>-7.68, 5.17</td>
<td>0.702</td>
<td>0.16</td>
<td>-6.23, 6.55</td>
<td>0.962</td>
</tr>
</tbody>
</table>

CI = confidence interval; MTX = methotrexate
Appendix H. Supplementary Primary Network Meta-Analyses

Overview of Content

This appendix contains the results of primary network meta-analyses (NWMA) based on studies with low or medium risk of bias but not shown in our main report because they rendered mostly inconclusive findings with wide confidence intervals. Specifically, these analyses evaluated Disease Activity Score (DAS) remission (Appendix Figure H-2). The network diagram for this outcome are presented in Appendix Figure H-1.

Additionally, we present full forest plots presenting our NWMA across all comparisons (and not within each comparison section) for every outcome of interest discussed in the main report: American College of Rheumatology response defined by 50 percent improvement (ACR50), radiographic joint damage, overall discontinuation, and discontinuation due to adverse events. These appear in Appendix Figure H-4, Appendix Figure H-6, Appendix Figure H-8, and Appendix Figure H-9, respectively, and network diagrams for these outcomes appear in Appendix Figure H-3, Appendix Figure H-5, and Appendix Figure H-7 (for both discontinuation outcomes), respectively.
Appendix Table H-1 lists the specific studies with low or medium risk of bias and reporting on outcomes of interest for our NWMA. These outcomes include DAS remission (n=10), ACR50 (n=11), radiographic joint damage (n=10), overall discontinuation (n=10), and discontinuation due to adverse events (n=12).
### Appendix Table H-1. Studies included in any KQ1 or KQ3 primary network meta-analyses

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>Study Name</th>
<th>DAS Remission&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Overall D/C&lt;sup&gt;a&lt;/sup&gt;</th>
<th>D/C due to AEs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ACR50&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Radiographic joint damage&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA+MTX vs. MTX</td>
<td>AGREE, 2009,31 2011,129,130 2015&lt;sup&gt;131&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ABA+MTX vs. ABA vs. MTX</td>
<td>AVERT, 2015&lt;sup&gt;17&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CZP+MTX vs. MTX</td>
<td>C-EARLY, 2017&lt;sup&gt;38, 39&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ETN vs. MTX</td>
<td>Enbrel ERA, 2000,14 2002,110 2005,164 2006&lt;sup&gt;111&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IFX+MTX vs. Methyl-PNL+MTX vs. MTX</td>
<td>Durez et al., 2007&lt;sup&gt;18&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IFX+MTX vs. MTX</td>
<td>Quinn et al., 2005&lt;sup&gt;31&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SSZ+MTX vs. SSZ vs. MTX</td>
<td>Dougados et al., 1999&lt;sup&gt;21&lt;/sup&gt; Maillefert et al., 2003&lt;sup&gt;104&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SSZ+MTX vs. SSZ vs. MTX</td>
<td>Haagsma et al., 1997&lt;sup&gt;23&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCZ+MTX vs. TCZ vs. MTX</td>
<td>FUNCTION, 2016,32 2017&lt;sup&gt;134&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCZ+MTX vs. TCZ vs. MTX</td>
<td>U-Act-Early, 2016&lt;sup&gt;13&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

ABA = abatacept; ACR50 = American College of Rheumatology 50% improvement; ADA = adalimumab; AE = adverse event; AGREE = Abatacept trial to Gauge Remission and joint damage progression in methotrexate-naïve patients with Early Erosive rheumatoid arthritis; ASPIRE = Active-controlled Study of Patients receiving Infliximab for the treatment of Rheumatoid arthritis of Early onset trial; AVERT = Assessing Very Early Rheumatoid arthritis Treatment trial; C-EARLY = trial whose acronym not described; C-OPERA = Certolizumab-Optimal Prevention of joint damage for Early RA trial; COMET = Combination of Methotrexate and Etanercept in Active Early Rheumatoid Arthritis trial; CZP = certolizumab pegol; D/C = discontinuation; DAS = Disease Activity Score; Enbrel ERA = Enbrel Early RA trial; ETN = etanercept; FUNCTION = trial whose acronym not described; IFX = infliximab; Methyl-PNL = methylprednisolone; MTX = methotrexate; NA = not applicable; NWMA = network meta-analysis; PREMIER = trial whose acronym not described; RA = rheumatoid arthritis; ROB = risk of bias; SSZ = sulfasalazine; TCZ = tocilizumab; U-Act-Early = Trial whose acronym not described; vs. = versus

<sup>a</sup> All data used in NWMA were measured at the 1-year follow-up timepoint.
Network Diagrams and Forest Plots

Appendix Figure H-1. Network diagram for network meta-analysis: Remission according to Disease Activity Score

MTX = methotrexate; N = number of patients
Appendix Figure H-2. Forest plots for network meta-analysis: Remission according to Disease Activity Score

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure H-3. Network diagram for network meta-analysis: ACR50 response

ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; N = number of patients
Appendix Figure H-4. Forest plots for network meta-analysis of ACR50 response

ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure H-5. Network diagram for network meta-analysis: Change from baseline in radiographic joint damage score

MTX = methotrexate; N = number of patients
Appendix Figure H-6. Forest plots for network meta-analysis: Change from baseline in radiographic joint damage score

MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. Infliximab+MTX</td>
<td>0.34 (0.10, 0.57)</td>
</tr>
<tr>
<td>Adalimumab vs. Infliximab+MTX</td>
<td>-0.18 (-0.42, 0.05)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Infliximab+MTX</td>
<td>-0.57 (-0.80, -0.33)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Infliximab+MTX</td>
<td>0.04 (-0.18, 0.26)</td>
</tr>
<tr>
<td>Etanercept vs. Infliximab+MTX</td>
<td>0.31 (0.06, 0.55)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Infliximab+MTX</td>
<td>-0.38 (-0.62, -0.15)</td>
</tr>
<tr>
<td>MTX vs. Infliximab+MTX</td>
<td>0.42 (0.27, 0.58)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Infliximab+MTX</td>
<td>0.45 (0.09, 0.82)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Infliximab+MTX</td>
<td>0.19 (-0.17, 0.56)</td>
</tr>
<tr>
<td>Tocilizumab vs. Infliximab+MTX</td>
<td>0.20 (-0.01, 0.40)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Infliximab+MTX</td>
<td>0.16 (-0.04, 0.37)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. MTX</td>
<td>-0.09 (-0.26, 0.09)</td>
</tr>
<tr>
<td>Adalimumab vs. MTX</td>
<td>-0.61 (-0.78, -0.43)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. MTX</td>
<td>-0.99 (-1.17, -0.81)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. MTX</td>
<td>-0.38 (-0.53, -0.23)</td>
</tr>
<tr>
<td>Etanercept vs. MTX</td>
<td>-0.12 (-0.31, 0.07)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. MTX</td>
<td>-0.81 (-0.98, -0.63)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. MTX</td>
<td>-0.42 (-0.58, -0.27)</td>
</tr>
<tr>
<td>Sulfasalazine vs. MTX</td>
<td>0.03 (-0.30, 0.37)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. MTX</td>
<td>-0.23 (-0.56, 0.10)</td>
</tr>
<tr>
<td>Tocilizumab vs. MTX</td>
<td>-0.23 (-0.37, -0.09)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. MTX</td>
<td>-0.26 (-0.40, -0.12)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Sulfasalazine</td>
<td>-0.12 (-0.49, 0.26)</td>
</tr>
<tr>
<td>Adalimumab vs. Sulfasalazine</td>
<td>-0.64 (-1.01, -0.26)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine</td>
<td>-1.02 (-1.40, -0.64)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine</td>
<td>-0.41 (-0.78, -0.04)</td>
</tr>
<tr>
<td>Etanercept vs. Sulfasalazine</td>
<td>-0.15 (-0.53, 0.24)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine</td>
<td>-0.84 (-1.21, -0.46)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine</td>
<td>-0.45 (-0.62, -0.09)</td>
</tr>
<tr>
<td>MTX vs. Sulfasalazine</td>
<td>-0.03 (-0.37, 0.30)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Sulfasalazine</td>
<td>-0.26 (-0.60, 0.07)</td>
</tr>
<tr>
<td>Tocilizumab vs. Sulfasalazine</td>
<td>-0.26 (-0.62, 0.10)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Sulfasalazine</td>
<td>-0.29 (-0.65, 0.07)</td>
</tr>
</tbody>
</table>
MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure H-7. Network diagram for network meta-analysis: All discontinuations and discontinuations due to adverse events

MTX = methotrexate; N = number of patients
Appendix Figure H-8. Forest plots for network meta-analysis: All discontinuations

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
<table>
<thead>
<tr>
<th>Study</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept vs. Tocilizumab+MTX</td>
<td>1.27 (0.71, 2.28)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Tocilizumab+MTX</td>
<td>0.64 (0.51, 1.38)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Tocilizumab+MTX</td>
<td>0.64 (0.39, 1.07)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Tocilizumab+MTX</td>
<td>0.63 (0.44, 0.90)</td>
</tr>
<tr>
<td>Etanercept vs. Tocilizumab+MTX</td>
<td>0.49 (0.30, 0.80)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Tocilizumab+MTX</td>
<td>0.64 (0.42, 1.00)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Tocilizumab+MTX</td>
<td>0.86 (0.58, 1.28)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Tocilizumab+MTX</td>
<td>0.56 (0.06, 5.01)</td>
</tr>
<tr>
<td>MTX vs. Tocilizumab+MTX</td>
<td>0.98 (0.74, 1.30)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Tocilizumab+MTX</td>
<td>1.80 (0.98, 3.28)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Tocilizumab+MTX</td>
<td>1.38 (0.75, 2.55)</td>
</tr>
<tr>
<td>Tocilizumab vs. Tocilizumab+MTX</td>
<td>0.84 (0.61, 1.14)</td>
</tr>
</tbody>
</table>

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure H-9. Forest plots for network meta-analysis: Discontinuations due to adverse events

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. Abatacept</td>
<td>0.34 (0.18, 0.64)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept</td>
<td>0.67 (0.19, 2.43)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept</td>
<td>0.75 (0.39, 1.45)</td>
</tr>
<tr>
<td>Etanercept vs. Abatacept</td>
<td>0.35 (0.17, 0.71)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept</td>
<td>0.56 (0.29, 1.08)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept</td>
<td>2.10 (0.93, 4.71)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Abatacept</td>
<td>0.69 (0.03, 15.41)</td>
</tr>
<tr>
<td>MTX vs. Abatacept</td>
<td>0.69 (0.44, 1.10)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Abatacept</td>
<td>1.42 (0.58, 3.45)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Abatacept</td>
<td>0.96 (0.38, 2.42)</td>
</tr>
<tr>
<td>Tocilizumab vs. Abatacept</td>
<td>1.16 (0.58, 2.31)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Abatacept</td>
<td>1.44 (0.75, 2.76)</td>
</tr>
<tr>
<td>Abatacept vs. Abatacept+MTX</td>
<td>2.95 (1.57, 5.54)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept+MTX</td>
<td>1.98 (0.53, 7.43)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept+MTX</td>
<td>2.21 (1.07, 4.57)</td>
</tr>
<tr>
<td>Etanercept vs. Abatacept+MTX</td>
<td>1.02 (0.47, 2.23)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept+MTX</td>
<td>1.64 (0.79, 3.40)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept+MTX</td>
<td>6.18 (2.59, 14.72)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Abatacept+MTX</td>
<td>2.02 (0.09, 46.13)</td>
</tr>
<tr>
<td>MTX vs. Abatacept+MTX</td>
<td>2.04 (1.17, 3.56)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Abatacept+MTX</td>
<td>4.17 (1.63, 10.71)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Abatacept+MTX</td>
<td>2.84 (1.07, 7.50)</td>
</tr>
<tr>
<td>Tocilizumab vs. Abatacept+MTX</td>
<td>3.42 (1.61, 7.28)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Abatacept+MTX</td>
<td>4.25 (2.07, 8.72)</td>
</tr>
<tr>
<td>Abatacept vs. Adalimumab+MTX</td>
<td>1.49 (0.41, 5.36)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab+MTX</td>
<td>0.50 (0.13, 1.89)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>1.11 (0.31, 4.02)</td>
</tr>
<tr>
<td>Etanercept vs. Adalimumab+MTX</td>
<td>0.51 (0.14, 1.92)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>0.83 (0.23, 2.99)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>3.11 (0.79, 12.25)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Adalimumab+MTX</td>
<td>1.02 (0.04, 27.69)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab+MTX</td>
<td>1.03 (0.31, 3.40)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Adalimumab+MTX</td>
<td>2.10 (0.51, 8.69)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Adalimumab+MTX</td>
<td>1.43 (0.34, 6.02)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab+MTX</td>
<td>1.73 (0.47, 6.33)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab+MTX</td>
<td>2.14 (0.60, 7.71)</td>
</tr>
</tbody>
</table>

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept vs. Sulfasalazine</td>
<td>0.71 (0.29, 1.72)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Sulfasalazine</td>
<td>0.24 (0.09, 0.61)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine</td>
<td>0.48 (0.12, 1.96)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine</td>
<td>0.53 (0.22, 1.29)</td>
</tr>
<tr>
<td>Etanercept vs. Sulfasalazine</td>
<td>0.24 (0.10, 0.62)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine</td>
<td>0.39 (0.16, 0.96)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine</td>
<td>1.48 (0.54, 4.06)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Sulfasalazine</td>
<td>0.48 (0.02, 11.53)</td>
</tr>
<tr>
<td>MTX vs. Sulfasalazine</td>
<td>0.49 (0.23, 1.04)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Sulfasalazine</td>
<td>0.68 (0.36, 1.29)</td>
</tr>
<tr>
<td>Tocilizumab vs. Sulfasalazine</td>
<td>0.82 (0.33, 2.05)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Sulfasalazine</td>
<td>1.02 (0.42, 2.47)</td>
</tr>
<tr>
<td>Abatacept vs. Sulfasalazine+MTX</td>
<td>1.04 (0.41, 2.61)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Sulfasalazine+MTX</td>
<td>0.35 (0.13, 0.93)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine+MTX</td>
<td>0.70 (0.17, 2.94)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine+MTX</td>
<td>0.78 (0.31, 1.96)</td>
</tr>
<tr>
<td>Etanercept vs. Sulfasalazine+MTX</td>
<td>0.36 (0.14, 0.95)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine+MTX</td>
<td>0.58 (0.23, 1.46)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine+MTX</td>
<td>2.18 (0.77, 6.14)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Sulfasalazine+MTX</td>
<td>0.71 (0.03, 17.11)</td>
</tr>
<tr>
<td>MTX vs. Sulfasalazine+MTX</td>
<td>0.72 (0.32, 1.59)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Sulfasalazine+MTX</td>
<td>1.47 (0.78, 2.78)</td>
</tr>
<tr>
<td>Tocilizumab vs. Sulfasalazine+MTX</td>
<td>1.21 (0.47, 3.10)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Sulfasalazine+MTX</td>
<td>1.50 (0.60, 3.75)</td>
</tr>
<tr>
<td>Abatacept vs. Tocilizumab</td>
<td>0.86 (0.43, 1.71)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Tocilizumab</td>
<td>0.29 (0.14, 0.62)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Tocilizumab</td>
<td>0.58 (0.16, 2.13)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Tocilizumab</td>
<td>0.65 (0.32, 1.29)</td>
</tr>
<tr>
<td>Etanercept vs. Tocilizumab</td>
<td>0.30 (0.14, 0.63)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Tocilizumab</td>
<td>0.48 (0.24, 0.96)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Tocilizumab</td>
<td>1.80 (0.78, 4.17)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Tocilizumab</td>
<td>0.59 (0.03, 13.37)</td>
</tr>
<tr>
<td>MTX vs. Tocilizumab</td>
<td>0.60 (0.36, 0.99)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Tocilizumab</td>
<td>1.22 (0.49, 3.04)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Tocilizumab</td>
<td>0.83 (0.32, 2.13)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Tocilizumab</td>
<td>1.24 (0.86, 1.79)</td>
</tr>
</tbody>
</table>

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix I. Sensitivity Analyses for Network Meta-Analyses

Sensitivity Analyses for Network Meta-Analyses

We identified a total of 14 studies with a low or medium risk of bias for use in our main network meta-analyses (NWMA) comparing the efficacy of drug therapies for early rheumatoid arthritis. Those findings are presented in our main report.

An additional two studies provided data eligible for inclusion in these analyses but were rated as high risk of bias. We re-ran our NWMA including these studies for our sensitivity analyses. Estimates for the treatment comparisons were very similar to estimates from our main analyses excluding those studies. We present these findings below, first for our tests of consistency and then the network diagrams and forest plots depicting effect estimates for specific drug comparisons.

Tests of Consistency: Models Including High Risk of Bias Studies

To test for consistency, we compared consistency and inconsistency models. In addition, where there were closed loops in the network diagram with both direct and indirect evidence available, we examined differences in results between direct and indirect evidence using network sidesplits.

ACR50 Response

For the ACR50 outcome (see Appendix Table I-1), there was no significant difference in the consistency and inconsistency models ($\chi^2(3) = 0.48$, $p=0.922$). Results did not differ significantly between direct and indirect evidence for (1) Abatacept versus Abatacept plus Methotrexate (MTX) (coefficient [95% CI] = -0.09 [-0.69 to 0.52], $p=0.777$), (2) Adalimumab versus Adalimumab plus MTX (coefficient [95% CI] = 0.17 [-0.55 to 0.89], $p=0.644$), or (3) Infliximab plus MTX versus Methylprednisolone plus MTX (coefficient [95% CI] = -0.37 [-1.99 to 1.25], $p=0.653$).

Remission According to Disease Activity Score

For the DAS outcome (see Appendix Table I-2), there was no significant difference in the consistency and inconsistency models ($\chi^2(2) = 1.66$, $p=0.646$). Results did not differ significantly between direct and indirect evidence for (1) Abatacept versus Abatacept + MTX (coefficient (95% CI) = -0.60 (-2.09, 0.89), $p=0.428$), or (2) Adalimumab versus Adalimumab + MTX (coefficient (95% CI) = -0.44 (-2.56 to 1.68), $p=0.685$).
### Appendix Table I-1. Table with network sidesplits: ACR50 Response

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept + MTX</td>
<td>0.16</td>
<td>-0.07, 0.38</td>
<td>0.178</td>
<td>0.24</td>
<td>-0.33, 0.406</td>
<td>0.178</td>
<td>-0.09</td>
<td>-0.69, 0.777</td>
<td>0.52</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>Adalimumab + MTX</td>
<td>0.42</td>
<td>0.25, 0.59</td>
<td>&lt;0.001</td>
<td>0.25</td>
<td>-0.47, 0.503</td>
<td>0.17</td>
<td>-0.55</td>
<td>-0.55, 0.644</td>
<td>0.89</td>
</tr>
<tr>
<td>Inflimimab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>0.00</td>
<td>-0.51, 0.18</td>
<td>1.000</td>
<td>0.37</td>
<td>-1.17, 0.636</td>
<td>0.37</td>
<td>-1.99, 0.653</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACR50 = American College of Rheumatology 50% response; CI = confidence interval; MTX = methotrexate

### Appendix Table I-2. Table with network sidesplits: Remission according to Disease Activity Score

<table>
<thead>
<tr>
<th>Drug A</th>
<th>Drug B</th>
<th>Direct Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Indirect Coefficient</th>
<th>95% CI</th>
<th>p</th>
<th>Coefficient Difference</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Abatacept + MTX</td>
<td>0.35</td>
<td>-0.18, 0.88</td>
<td>0.192</td>
<td>0.95</td>
<td>-0.45, 0.184</td>
<td>0.184</td>
<td>-0.60</td>
<td>-2.09, 0.428</td>
<td>0.89</td>
</tr>
<tr>
<td>Inflimimab+MTX</td>
<td>Methylprednisolone+MTX</td>
<td>0.10</td>
<td>-0.56, 0.75</td>
<td>0.777</td>
<td>0.53</td>
<td>-1.47, 2.54</td>
<td>0.44</td>
<td>-2.56, 1.68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI = confidence interval; MTX = methotrexate
Network Diagrams and Forest Plots

Appendix Figure I-1. Network diagram for network meta-analysis (sensitivity analysis): ACR50 response

MTX = methotrexate; N = number of patients
Figure I-2 displays a forest plot for the sensitivity analysis of studies reporting data for ACR50 response rates, including studies with a high risk of bias. Study-level data used in this Figure are presented in Appendix C. We repeated the network meta-analyses (NWMA) including two
studies with high risk of bias as a sensitivity analysis: another eligible study of CZP plus MTX and another study of ADA plus MTX. This figure is described further in Appendix I as follows: “Estimates for the treatment comparisons were very similar to estimates from our main analyses excluding those studies.”

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept vs. Adalimumab+MTX</td>
<td>0.85 (0.65, 1.22)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab+MTX</td>
<td>1.01 (0.82, 1.24)</td>
</tr>
<tr>
<td>Adalimumab vs. Adalimumab+MTX</td>
<td>0.66 (0.56, 0.79)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>0.96 (0.80, 1.16)</td>
</tr>
<tr>
<td>Etanercept vs. Adalimumab+MTX</td>
<td>0.71 (0.56, 0.91)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>1.12 (0.90, 1.39)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>1.18 (0.94, 1.49)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Adalimumab+MTX</td>
<td>1.22 (0.73, 2.06)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab+MTX</td>
<td>0.75 (0.65, 0.87)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab+MTX</td>
<td>0.90 (0.73, 1.11)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab+MTX</td>
<td>0.97 (0.80, 1.18)</td>
</tr>
<tr>
<td>Abatacept vs. Certolizumab+MTX</td>
<td>0.88 (0.68, 1.14)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Certolizumab+MTX</td>
<td>1.05 (0.87, 1.25)</td>
</tr>
<tr>
<td>Adalimumab vs. Certolizumab+MTX</td>
<td>0.69 (0.55, 0.86)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Certolizumab+MTX</td>
<td>1.04 (0.86, 1.24)</td>
</tr>
<tr>
<td>Etanercept vs. Certolizumab+MTX</td>
<td>0.74 (0.59, 0.92)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Certolizumab+MTX</td>
<td>1.16 (0.96, 1.41)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Certolizumab+MTX</td>
<td>1.22 (0.99, 1.52)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Certolizumab+MTX</td>
<td>1.27 (0.76, 2.12)</td>
</tr>
<tr>
<td>MTX vs. Certolizumab+MTX</td>
<td>0.78 (0.70, 0.87)</td>
</tr>
<tr>
<td>Tocilizumab vs. Certolizumab+MTX</td>
<td>0.93 (0.77, 1.12)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Certolizumab+MTX</td>
<td>1.01 (0.84, 1.20)</td>
</tr>
<tr>
<td>Abatacept vs. Etanercept</td>
<td>1.20 (0.89, 1.62)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Etanercept</td>
<td>1.42 (1.11, 1.80)</td>
</tr>
<tr>
<td>Adalimumab vs. Etanercept</td>
<td>0.93 (0.71, 1.22)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Etanercept</td>
<td>1.40 (1.10, 1.79)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Etanercept</td>
<td>1.35 (1.08, 1.69)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Etanercept</td>
<td>1.57 (1.23, 2.02)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Etanercept</td>
<td>1.66 (1.27, 2.16)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Etanercept</td>
<td>1.72 (1.00, 2.94)</td>
</tr>
<tr>
<td>MTX vs. Etanercept</td>
<td>1.06 (0.87, 1.28)</td>
</tr>
<tr>
<td>Tocilizumab vs. Etanercept</td>
<td>1.26 (0.99, 1.61)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Etanercept</td>
<td>1.36 (1.08, 1.72)</td>
</tr>
</tbody>
</table>
Figure I-2 displays a forest plot for the sensitivity analysis of studies reporting data for ACR50 response rates, including studies with a high risk of bias. Study-level data used in this Figure are presented in Appendix C. We repeated the network meta-analyses (NWMA) including two studies with high risk of bias as a sensitivity analysis: another eligible study of CZP plus MTX and another study of ADA plus MTX. This figure is described further in Appendix I as follows: “Estimates for the treatment comparisons were very similar to estimates from our main analyses excluding those studies”.

---

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept vs. Etanercept+MTX</td>
<td>0.76 (0.58, 1.01)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Etanercept+MTX</td>
<td>0.90 (0.73, 1.11)</td>
</tr>
<tr>
<td>Adalimumab vs. Etanercept+MTX</td>
<td>0.59 (0.46, 0.76)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Etanercept+MTX</td>
<td>0.89 (0.72, 1.11)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Etanercept+MTX</td>
<td>0.86 (0.71, 1.04)</td>
</tr>
<tr>
<td>Etanercept vs. Etanercept+MTX</td>
<td>0.63 (0.49, 0.82)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Etanercept+MTX</td>
<td>1.05 (0.83, 1.34)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Etanercept+MTX</td>
<td>1.09 (0.65, 1.84)</td>
</tr>
<tr>
<td>MTX vs. Etanercept+MTX</td>
<td>0.67 (0.57, 0.79)</td>
</tr>
<tr>
<td>Tocilizumab vs. Etanercept+MTX</td>
<td>0.80 (0.65, 1.00)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Etanercept+MTX</td>
<td>0.86 (0.70, 1.06)</td>
</tr>
<tr>
<td>Abatacept vs. Infliximab+MTX</td>
<td>0.72 (0.54, 0.97)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Infliximab+MTX</td>
<td>0.85 (0.68, 1.08)</td>
</tr>
<tr>
<td>Adalimumab vs. Infliximab+MTX</td>
<td>0.56 (0.43, 0.73)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Infliximab+MTX</td>
<td>0.85 (0.67, 1.07)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Infliximab+MTX</td>
<td>0.82 (0.66, 1.01)</td>
</tr>
<tr>
<td>Etanercept vs. Infliximab+MTX</td>
<td>0.60 (0.46, 0.79)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Infliximab+MTX</td>
<td>0.95 (0.75, 1.21)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Infliximab+MTX</td>
<td>1.04 (0.64, 1.68)</td>
</tr>
<tr>
<td>MTX vs. Infliximab+MTX</td>
<td>0.64 (0.53, 0.77)</td>
</tr>
<tr>
<td>Tocilizumab vs. Infliximab+MTX</td>
<td>0.76 (0.60, 0.96)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Infliximab+MTX</td>
<td>0.82 (0.66, 1.03)</td>
</tr>
<tr>
<td>Abatacept vs. Methylprednisolone+MTX</td>
<td>0.70 (0.40, 1.21)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Methylprednisolone+MTX</td>
<td>0.82 (0.49, 1.39)</td>
</tr>
<tr>
<td>Adalimumab vs. Methylprednisolone+MTX</td>
<td>0.54 (0.32, 0.93)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Methylprednisolone+MTX</td>
<td>0.82 (0.49, 1.38)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Methylprednisolone+MTX</td>
<td>0.79 (0.47, 1.32)</td>
</tr>
<tr>
<td>Etanercept vs. Methylprednisolone+MTX</td>
<td>0.58 (0.34, 1.00)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Methylprednisolone+MTX</td>
<td>0.92 (0.54, 1.55)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Methylprednisolone+MTX</td>
<td>0.96 (0.60, 1.56)</td>
</tr>
<tr>
<td>MTX vs. Methylprednisolone+MTX</td>
<td>0.62 (0.37, 1.02)</td>
</tr>
<tr>
<td>Tocilizumab vs. Methylprednisolone+MTX</td>
<td>0.74 (0.44, 1.24)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Methylprednisolone+MTX</td>
<td>0.79 (0.47, 1.33)</td>
</tr>
</tbody>
</table>
ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure I-3. Forest plots for network meta-analysis (sensitivity analysis) of ACR50 response: Comparison of combined therapies to MTX only

ACR50 = American College of Rheumatology 50% improvement; MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure I-4. Network diagram for network meta-analysis (sensitivity analysis): Change from baseline in radiographic joint damage score

MTX = methotrexate; N = number of patients
Appendix Figure I-5. Forest plots for network meta-analysis (sensitivity analysis): Change from baseline in radiographic joint damage score

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adalimumab vs. Abatacept+MTX</td>
<td>-0.52 (-0.77, -0.28)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept+MTX</td>
<td>-0.90 (-1.15, -0.65)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept+MTX</td>
<td>-0.27 (-0.49, -0.06)</td>
</tr>
<tr>
<td>Etanercept vs. Abatacept+MTX</td>
<td>-0.03 (-0.29, 0.23)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept+MTX</td>
<td>-0.72 (-0.97, -0.47)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept+MTX</td>
<td>-0.34 (-0.57, -0.10)</td>
</tr>
<tr>
<td>MTX vs. Abatacept+MTX</td>
<td>0.09 (-0.09, 0.26)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Abatacept+MTX</td>
<td>0.12 (-0.26, 0.49)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Abatacept+MTX</td>
<td>-0.15 (-0.52, 0.23)</td>
</tr>
<tr>
<td>Tocilizumab vs. Abatacept+MTX</td>
<td>-0.14 (-0.36, 0.08)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Abatacept+MTX</td>
<td>-0.17 (-0.39, 0.05)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab</td>
<td>0.52 (0.28, 0.77)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Adalimumab</td>
<td>-0.38 (-0.55, -0.21)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab</td>
<td>0.25 (0.03, 0.46)</td>
</tr>
<tr>
<td>Etanercept vs. Adalimumab</td>
<td>0.49 (0.23, 0.75)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab</td>
<td>-0.20 (-0.44, 0.05)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab</td>
<td>0.18 (-0.05, 0.42)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab</td>
<td>0.61 (0.43, 0.78)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Adalimumab</td>
<td>0.64 (0.26, 1.01)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Adalimumab</td>
<td>0.38 (0.00, 0.75)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab</td>
<td>0.38 (0.16, 0.60)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab</td>
<td>0.35 (0.13, 0.57)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab+MTX</td>
<td>0.90 (0.65, 1.15)</td>
</tr>
<tr>
<td>Adalimumab vs. Adalimumab+MTX</td>
<td>0.38 (0.21, 0.55)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>0.63 (0.41, 0.85)</td>
</tr>
<tr>
<td>Etanercept vs. Adalimumab+MTX</td>
<td>0.87 (0.61, 1.13)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>0.18 (-0.07, 0.43)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>0.57 (0.33, 0.80)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab+MTX</td>
<td>0.99 (0.81, 1.17)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Adalimumab+MTX</td>
<td>1.02 (0.64, 1.40)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Adalimumab+MTX</td>
<td>0.76 (0.38, 1.13)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab+MTX</td>
<td>0.76 (0.54, 0.99)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab+MTX</td>
<td>0.73 (0.50, 0.96)</td>
</tr>
<tr>
<td>Treatment Comparison</td>
<td>SMD (95% CI)</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Sulfasalazine+MTX</td>
<td>0.15 (-0.23, 0.52)</td>
</tr>
<tr>
<td>Adalimumab vs. Sulfasalazine+MTX</td>
<td>-0.38 (-0.75, -0.00)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Sulfasalazine+MTX</td>
<td>-0.76 (-1.13, -0.38)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Sulfasalazine+MTX</td>
<td>-0.13 (-0.48, 0.23)</td>
</tr>
<tr>
<td>Etanercept vs. Sulfasalazine+MTX</td>
<td>0.11 (-0.27, 0.50)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Sulfasalazine+MTX</td>
<td>-0.57 (-0.95, -0.20)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Sulfasalazine+MTX</td>
<td>-0.19 (-0.56, 0.17)</td>
</tr>
<tr>
<td>MTX vs. Sulfasalazine+MTX</td>
<td>0.23 (-0.10, 0.56)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Sulfasalazine+MTX</td>
<td>0.26 (-0.07, 0.60)</td>
</tr>
<tr>
<td>Tocilizumab vs. Sulfasalazine+MTX</td>
<td>0.00 (-0.35, 0.36)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Sulfasalazine+MTX</td>
<td>-0.03 (-0.39, 0.33)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Tocilizumab</td>
<td>0.14 (-0.08, 0.36)</td>
</tr>
<tr>
<td>Adalimumab vs. Tocilizumab</td>
<td>-0.38 (-0.60, -0.16)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Tocilizumab</td>
<td>-0.76 (-0.99, -0.54)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Tocilizumab</td>
<td>-0.13 (-0.32, 0.06)</td>
</tr>
<tr>
<td>Etanercept vs. Tocilizumab</td>
<td>0.11 (-0.13, 0.35)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Tocilizumab</td>
<td>-0.58 (-0.80, -0.35)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Tocilizumab</td>
<td>-0.20 (-0.40, 0.01)</td>
</tr>
<tr>
<td>MTX vs. Tocilizumab</td>
<td>0.23 (0.09, 0.37)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Tocilizumab</td>
<td>0.26 (-0.10, 0.62)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Tocilizumab</td>
<td>-0.00 (-0.36, 0.35)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Tocilizumab</td>
<td>-0.03 (-0.17, 0.11)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Tocilizumab+MTX</td>
<td>0.17 (-0.05, 0.39)</td>
</tr>
<tr>
<td>Adalimumab vs. Tocilizumab+MTX</td>
<td>-0.35 (-0.57, -0.13)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Tocilizumab+MTX</td>
<td>-0.73 (-0.96, -0.50)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Tocilizumab+MTX</td>
<td>-0.10 (-0.29, 0.09)</td>
</tr>
<tr>
<td>Etanercept vs. Tocilizumab+MTX</td>
<td>0.14 (-0.10, 0.38)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Tocilizumab+MTX</td>
<td>-0.55 (-0.77, -0.32)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Tocilizumab+MTX</td>
<td>-0.16 (-0.37, 0.04)</td>
</tr>
<tr>
<td>MTX vs. Tocilizumab+MTX</td>
<td>0.26 (0.12, 0.40)</td>
</tr>
<tr>
<td>Sulfasalazine vs. Tocilizumab+MTX</td>
<td>0.29 (-0.07, 0.65)</td>
</tr>
<tr>
<td>Sulfasalazine+MTX vs. Tocilizumab+MTX</td>
<td>0.03 (-0.33, 0.39)</td>
</tr>
<tr>
<td>Tocilizumab vs. Tocilizumab+MTX</td>
<td>0.03 (-0.11, 0.17)</td>
</tr>
</tbody>
</table>

MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval.
Appendix Figure I-6. Forest plots for network meta-analysis (sensitivity analysis) of change from baseline in radiographic joint damage score: Comparison of combined therapies to MTX only

MTX = methotrexate; SMD = standardized mean difference; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure I-7. Network diagram for network meta-analysis (sensitivity analysis): Remission according to Disease Activity Score

MTX = methotrexate; N = number of patients
Appendix Figure I-8. Forest plots for network meta-analysis (sensitivity analysis): Remission according to Disease Activity Score

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept+MTX vs. Abatacept</td>
<td>1.52 (0.94, 2.47)</td>
</tr>
<tr>
<td>Adalimumab vs. Abatacept</td>
<td>0.99 (0.49, 2.00)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept</td>
<td>1.70 (0.91, 3.19)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept</td>
<td>1.59 (0.86, 2.91)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept</td>
<td>1.77 (0.88, 3.56)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept</td>
<td>1.58 (0.81, 3.10)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Abatacept</td>
<td>1.82 (0.79, 4.22)</td>
</tr>
<tr>
<td>MTX vs. Abatacept</td>
<td>1.00 (0.61, 1.64)</td>
</tr>
<tr>
<td>Tocilizumab vs. Abatacept</td>
<td>1.48 (0.81, 2.71)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Abatacept</td>
<td>1.50 (0.82, 2.74)</td>
</tr>
<tr>
<td>Abatacept vs. Abatacept+MTX</td>
<td>0.66 (0.40, 1.07)</td>
</tr>
<tr>
<td>Adalimumab vs. Abatacept+MTX</td>
<td>0.65 (0.35, 1.21)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Abatacept+MTX</td>
<td>1.12 (0.66, 1.90)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Abatacept+MTX</td>
<td>1.04 (0.63, 1.73)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Abatacept+MTX</td>
<td>1.16 (0.63, 2.14)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Abatacept+MTX</td>
<td>1.04 (0.58, 1.86)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Abatacept+MTX</td>
<td>1.20 (0.55, 2.58)</td>
</tr>
<tr>
<td>MTX vs. Abatacept+MTX</td>
<td>0.66 (0.46, 0.94)</td>
</tr>
<tr>
<td>Tocilizumab vs. Abatacept+MTX</td>
<td>0.97 (0.59, 1.60)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Abatacept+MTX</td>
<td>0.99 (0.60, 1.62)</td>
</tr>
<tr>
<td>Abatacept vs. Adalimumab</td>
<td>1.01 (0.50, 2.06)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab</td>
<td>1.54 (0.83, 2.88)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Adalimumab</td>
<td>1.73 (1.06, 2.81)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab</td>
<td>1.61 (0.87, 2.99)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab</td>
<td>1.79 (0.88, 3.64)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab</td>
<td>1.60 (0.81, 3.17)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Adalimumab</td>
<td>1.85 (0.79, 4.31)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab</td>
<td>1.01 (0.61, 1.68)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab</td>
<td>1.50 (0.81, 2.78)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab</td>
<td>1.52 (0.83, 2.81)</td>
</tr>
<tr>
<td>Comparison</td>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Abatacept vs. Adalimumab+MTX</td>
<td>0.59 (0.31, 1.10)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Adalimumab+MTX</td>
<td>0.89 (0.53, 1.52)</td>
</tr>
<tr>
<td>Adalimumab vs. Adalimumab+MTX</td>
<td>0.58 (0.36, 0.94)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Adalimumab+MTX</td>
<td>0.93 (0.55, 1.58)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Adalimumab+MTX</td>
<td>1.04 (0.55, 1.95)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Adalimumab+MTX</td>
<td>0.93 (0.51, 1.69)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Adalimumab+MTX</td>
<td>1.07 (0.49, 2.34)</td>
</tr>
<tr>
<td>MTX vs. Adalimumab+MTX</td>
<td>0.59 (0.40, 0.87)</td>
</tr>
<tr>
<td>Tocilizumab vs. Adalimumab+MTX</td>
<td>0.87 (0.52, 1.47)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Adalimumab+MTX</td>
<td>0.88 (0.52, 1.48)</td>
</tr>
<tr>
<td>Abatacept vs. Certolizumab+MTX</td>
<td>0.63 (0.34, 1.16)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Certolizumab+MTX</td>
<td>0.96 (0.58, 1.59)</td>
</tr>
<tr>
<td>Adalimumab vs. Certolizumab+MTX</td>
<td>0.62 (0.33, 1.16)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Certolizumab+MTX</td>
<td>1.07 (0.63, 1.82)</td>
</tr>
<tr>
<td>Etanercept+MTX vs. Certolizumab+MTX</td>
<td>1.11 (0.61, 2.05)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Certolizumab+MTX</td>
<td>1.00 (0.56, 1.78)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Certolizumab+MTX</td>
<td>1.15 (0.53, 2.47)</td>
</tr>
<tr>
<td>MTX vs. Certolizumab+MTX</td>
<td>0.63 (0.44, 0.90)</td>
</tr>
<tr>
<td>Tocilizumab vs. Certolizumab+MTX</td>
<td>0.94 (0.57, 1.53)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Certolizumab+MTX</td>
<td>0.95 (0.58, 1.55)</td>
</tr>
<tr>
<td>Abatacept vs. Etanercept+MTX</td>
<td>0.57 (0.28, 1.14)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. Etanercept+MTX</td>
<td>0.86 (0.47, 1.59)</td>
</tr>
<tr>
<td>Adalimumab vs. Etanercept+MTX</td>
<td>0.56 (0.27, 1.13)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. Etanercept+MTX</td>
<td>0.96 (0.51, 1.81)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. Etanercept+MTX</td>
<td>0.90 (0.49, 1.65)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. Etanercept+MTX</td>
<td>0.90 (0.46, 1.75)</td>
</tr>
<tr>
<td>Methylprednisolone+MTX vs. Etanercept+MTX</td>
<td>1.03 (0.44, 2.39)</td>
</tr>
<tr>
<td>MTX vs. Etanercept+MTX</td>
<td>0.57 (0.34, 0.93)</td>
</tr>
<tr>
<td>Tocilizumab vs. Etanercept+MTX</td>
<td>0.84 (0.46, 1.54)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. Etanercept+MTX</td>
<td>0.85 (0.46, 1.55)</td>
</tr>
</tbody>
</table>
MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix Figure I-9. Forest plots for network meta-analysis (sensitivity analysis) of remission according to Disease Activity Score: Comparison of combined therapies to MTX only

<table>
<thead>
<tr>
<th>Methylprednisolone+MTX vs. MTX</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etanercept+MTX vs. MTX</td>
<td>1.77 (1.07, 2.92)</td>
</tr>
<tr>
<td>Adalimumab+MTX vs. MTX</td>
<td>1.70 (1.15, 2.51)</td>
</tr>
<tr>
<td>Certolizumab+MTX vs. MTX</td>
<td>1.58 (1.12, 2.27)</td>
</tr>
<tr>
<td>Infliximab+MTX vs. MTX</td>
<td>1.58 (1.01, 2.48)</td>
</tr>
<tr>
<td>Abatacept+MTX vs. MTX</td>
<td>1.52 (1.06, 2.18)</td>
</tr>
<tr>
<td>Tocilizumab+MTX vs. MTX</td>
<td>1.51 (1.06, 2.12)</td>
</tr>
</tbody>
</table>

MTX = methotrexate; RR = relative risk; vs. = versus; 95% CI = 95% confidence interval
Appendix J. Expert Guidance and Review

Stakeholder Input in Formulating the Research Protocol

Stakeholders, including Key Informants and Technical Experts, participated in a virtual workshop by PCORI in December 2016 to help formulate the research protocol. Details on the virtual workshop, including a list of participants, can be found at https://www.pcori.org/events/2016/updating-systematic-reviews-pcori-virtual-multi-stakeholder-workshop-drug-therapy.

Key Informants in the workshop included end users of research, such as patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Technical Experts in the workshop included multidisciplinary groups of clinical, content, and methodological experts who provided input in defining populations, interventions, comparisons, and outcomes, and identified particular studies or databases to search. They were selected to provide broad expertise and perspectives specific to rheumatoid arthritis (RA).

During the virtual workshop, stakeholders reviewed scoping for the updated review, prioritized Key Questions, and discussed where the evidence base has accumulated since the prior review, as well as emerging issues in RA. Based upon findings from the workshop, the RA protocol was developed by the EPC with guidance from PCORI and AHRQ.

Key Informants and Technical Experts did not do analysis of any kind or contribute to the writing of this draft report. They will be given the opportunity to review the report through the peer or public review mechanisms.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers follows:

Joan M. Bathon, M.D.
Director, Division of Rheumatology
NewYork-Presbyterian Hospital/Columbia University Medical Center
New York, NY

Rongwei (Rochelle) Fu, Ph.D.
Director, Biostatistics Education Program – School of Public Health
Oregon Health Sciences University
Portland, OR
Suzanne Schrandt, J.D.
Director, Patient Engagement
Arthritis Foundation
Atlanta, GA

Jasvinder Singh, M.D., M.P.H.
Professor of Medicine and Epidemiology
University of Alabama at Birmingham
Birmingham, AL
### Appendix K. PCORI Methodology Standards Checklist: SER Update

<table>
<thead>
<tr>
<th>Contract No.</th>
<th>HHSA290201500011I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Order No.</td>
<td>10</td>
</tr>
<tr>
<td>EPC</td>
<td>RTI-UNC</td>
</tr>
<tr>
<td>Project Title</td>
<td>Drug Therapy for Early Rheumatoid Arthritis in Adults</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Category</th>
<th>Standard</th>
<th>Is this standard applicable to this SER update?</th>
<th>List sections and pages of the SER report where you address this standard</th>
<th>If applicable, describe how and why the SER update deviated from this standard?</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ-1</td>
<td>Identify Gaps in Evidence</td>
<td>Yes</td>
<td>Intro: pages 3-4</td>
<td></td>
</tr>
<tr>
<td>RQ-2</td>
<td>Develop a Formal Study Protocol</td>
<td>Yes</td>
<td>Published Protocol on AHRQ EHC website</td>
<td></td>
</tr>
<tr>
<td>RQ-3</td>
<td>Identify Specific Populations and Health Decision(s) Affected by the Research</td>
<td>Yes</td>
<td>Intro: pages 4-6; Methods: page 7</td>
<td></td>
</tr>
<tr>
<td>RQ-4</td>
<td>Identify and Assess Participant Subgroups</td>
<td>Yes</td>
<td>Intro: page 5; Methods: page 7</td>
<td></td>
</tr>
<tr>
<td>RQ-5</td>
<td>Select Appropriate Interventions and Comparators</td>
<td>Yes</td>
<td>Intro: pages 2-6; Methods: pages 7-8</td>
<td></td>
</tr>
<tr>
<td>RQ-6</td>
<td>Measure Outcomes that People Representing the Population of Interest Notice and Care About</td>
<td>Yes</td>
<td>Intro: pages 4-6; Methods: pages 7-8, 10-11</td>
<td></td>
</tr>
<tr>
<td>Contract No.</td>
<td>HHSA290201500011I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Order No.</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPC</td>
<td>RTI-UNC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Title</td>
<td>Drug Therapy for Early Rheumatoid Arthritis in Adults</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard Category</strong></td>
<td><strong>Abbrev.</strong></td>
<td><strong>Standard</strong></td>
<td><strong>Is this standard applicable to this SER update?</strong></td>
<td><strong>List sections and pages of the SER report where you address this standard</strong></td>
</tr>
<tr>
<td><strong>Standards Associated with Patient-Centeredness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC-1</td>
<td>Engage people representing the population of interest and other relevant stakeholders in ways that are appropriate and necessary in a given research context.</td>
<td>Yes</td>
<td>Front Matter: page iii; page 7 Appendix J</td>
<td>Refers to the PCORI Stakeholder Call in December 2016 that gathered stakeholder groups and technical experts to define the scope of this review.</td>
</tr>
<tr>
<td>PC-2</td>
<td>Identify, Select, Recruit, and Retain Study Participants Representative of the Spectrum of the Population of Interest and Ensure that Data Are Collected Thoroughly and Systematically from All Study Participants</td>
<td>N/A</td>
<td></td>
<td>Systematic review with no primary data collection.</td>
</tr>
<tr>
<td>PC-3</td>
<td>Use Patient-Reported Outcomes When Patients or People at Risk of a Condition Are the Best Source of Information</td>
<td>N/A</td>
<td></td>
<td>Systematic review with no primary data collection. However, we used patient-centered outcomes data for KQ2 whenever our included studies reported them.</td>
</tr>
<tr>
<td>PC-4</td>
<td>Support dissemination and implementation of study results</td>
<td>N/A</td>
<td></td>
<td>Systematic review with no primary data collection.</td>
</tr>
<tr>
<td>IR-1</td>
<td>Assess Data Source Adequacy</td>
<td>Yes</td>
<td>Methods: pages 8-13</td>
<td></td>
</tr>
<tr>
<td>Contract No.</td>
<td>HHSA290201500011I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Order No.</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPC</td>
<td>RTI-UNC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Title</td>
<td>Drug Therapy for Early Rheumatoid Arthritis in Adults</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Category</th>
<th>Abbrev.</th>
<th>Standard</th>
<th>Is this standard applicable to this SER update?</th>
<th>List sections and pages of the SER report where you address this standard</th>
<th>If applicable, describe how and why the SER update deviated from this standard?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standards for Data Integrity and Rigorous Analyses</td>
<td>IR-2</td>
<td>Describe Data Linkage Plans, if Applicable</td>
<td>N/A</td>
<td></td>
<td>No data linkage required.</td>
</tr>
<tr>
<td></td>
<td>IR-3</td>
<td>A priori, Specify Plans for Data Analysis that Correspond to Major Aims</td>
<td>Yes</td>
<td>Methods: pages 12-13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IR-4</td>
<td>Document Validated Scales and Tests</td>
<td>Yes</td>
<td>Methods: pages 10-11; Appendix F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IR-5</td>
<td>Use Sensitivity Analyses to Determine the Impact of Key Assumptions</td>
<td>Yes</td>
<td>Methods (high ROB): page 12; Results: page 41; Discussion: page 115; Appendix I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IR-6</td>
<td>Provide Sufficient Information in Reports to Allow for Assessments of the Study’s Internal and External Validity</td>
<td>Yes</td>
<td>Methods: pages 11-12, 14; Discussion: pages 125-126; Appendix D</td>
<td></td>
</tr>
<tr>
<td>Standards for Preventing and Handling Missing Data</td>
<td>MD-1</td>
<td>Describe in Protocol Methods to Prevent and Monitor Missing Data</td>
<td>Yes</td>
<td>Methods (handsearching, gray literature, SEADs): pages 9, 11-12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MD-2</td>
<td>Describe Statistical Methods to Handle Missing Data in Protocol</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Standard Category</td>
<td>Abbrev.</td>
<td>Standard</td>
<td>Is this standard applicable to this SER update?</td>
<td>List sections and pages of the SER report where you address this standard</td>
<td>If applicable, describe how and why the SER update deviated from this standard?</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>----------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>MD-3</td>
<td>Use Validated Methods to Deal with Missing Data that Properly Account for Statistical Uncertainty Due to Missingness</td>
<td>Yes</td>
<td>Methods (imputation plan for missing network meta-analysis data): page 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MD-4</td>
<td>Record and Report All Reasons for Dropout and Missing Data, and Account for All Patients in Reports</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>MD-5</td>
<td>Examine Sensitivity of Inferences to Missing Data Methods and Assumptions, and Incorporate into Interpretation</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Standards for Heterogeneity of Treatment Effect (HTE)</td>
<td>HT-1</td>
<td>State the Goals of HTE Analyses</td>
<td>Yes</td>
<td>Methods: page 12; Discussion: page 128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HT-2</td>
<td>For all HTE Analyses, Pre-specify the analysis plan; for Hypothesis driven HTE Analyses, Pre-specify Hypotheses and supporting evidence base</td>
<td>Yes</td>
<td>Methods: page 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HT-3</td>
<td>All HTE claims must be based on appropriate statistical contrasts among groups being compared, such as interaction tests or estimates of differences in treatment effect</td>
<td>Yes</td>
<td>Discussion (Limitations): pages 127-128</td>
<td></td>
</tr>
<tr>
<td>Standard Category</td>
<td>Abbrev.</td>
<td>Standard</td>
<td>Is this standard applicable to this SER update?</td>
<td>List sections and pages of the SER report where you address this standard</td>
<td>If applicable, describe how and why the SER update deviated from this standard?</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Standards for Specific Study Designs and Methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standards for Data Registries</td>
<td>DR-1</td>
<td>Requirements for the Design and Features of Registries</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DR-2</td>
<td>Standards for Selection and Use of Registries</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DR-3</td>
<td>Robust Analysis of Confounding Factors</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Standards for Data Networks as Research-Facilitating Structures</td>
<td>DN-1</td>
<td>Requirements for the Design and Features of Data Networks</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DN-2</td>
<td>Standards for Selection and Use of Data Networks</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Causal Inference Standards</td>
<td>CI-1</td>
<td>Define Analysis Population Using Covariate Histories</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Standard Category</td>
<td>Abbrev.</td>
<td>Standard</td>
<td>Is this standard applicable to this SER update?</td>
<td>List sections and pages of the SER report where you address this standard</td>
<td>If applicable, describe how and why the SER update deviated from this standard?</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CI-2</td>
<td></td>
<td>Describe Population that Gave Rise to the Effect Estimate(s)</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>CI-3</td>
<td></td>
<td>Precisely Define the Timing of the Outcome Assessment Relative to the Initiation and Duration of Exposure</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>CI-4</td>
<td></td>
<td>Measure Confounders before Start of Exposure. Report data on confounders with study results</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>CI-5</td>
<td></td>
<td>Report the assumptions underlying the construction of Propensity Scores and the comparability of the resulting groups in terms of the balance of covariates and overlap</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>CI-6</td>
<td></td>
<td>Assess the Validity of the Instrumental Variable (i.e. how the assumption are met) and report the balance of covariates in the groups created by the IV for all IV analyses</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>AT-1</td>
<td></td>
<td>Specify Planned Adaptations and Primary Analysis</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>AT-2</td>
<td></td>
<td>Evaluate Statistical Properties of Adaptive Design</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Contract No.</td>
<td>HHSA290201500011I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Order No.</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPC</td>
<td>RTI-UNC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Title</td>
<td>Drug Therapy for Early Rheumatoid Arthritis in Adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard Category</th>
<th>Abbrev.</th>
<th>Standard</th>
<th>Is this standard applicable to this SER update?</th>
<th>List sections and pages of the SER report where you address this standard</th>
<th>If applicable, describe how and why the SER update deviated from this standard?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AT-3</td>
<td>Specify Structure and Analysis Plan for Bayesian Adaptive Randomized Clinical Trial Designs</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>AT-4</td>
<td>Ensure Clinical Trial Infrastructure Is Adequate to Support Planned Adaptation(s)</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>AT-5</td>
<td>Use the CONSORT statement, with Modifications, to Report Adaptive Randomized Clinical Trials</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DT-1</td>
<td>Specify Clinical Context and Key Elements of Diagnostic Test Study Design</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DT-2</td>
<td>Study Design Should be Informed by Investigations of the Clinical Context of Testing</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DT-3</td>
<td>Assess the Effect of Factors Known to Affect Diagnostic Performance and Outcomes</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td></td>
<td>DT-4</td>
<td>Structured Reporting of Diagnostic Comparative Effectiveness Study Results</td>
<td>N/A</td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Standard Category</td>
<td>Abbrev.</td>
<td>Standard</td>
<td>Is this standard applicable to this SER update?</td>
<td>List sections and pages of the SER report where you address this standard</td>
<td>If applicable, describe how and why the SER update deviated from this standard?</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>----------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Focus studies of diagnostic tests on patient centered outcomes, using rigorous study designs with preference for randomized controlled trials</td>
<td>DT-5</td>
<td>N/A</td>
<td></td>
<td></td>
<td>Standard does not apply.</td>
</tr>
<tr>
<td>Adopt the Institute of Medicine (IOM) standards for systematic reviews of comparative effectiveness research, with some qualifications.</td>
<td>SR-1</td>
<td>Yes</td>
<td>Entire report (all pages)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


114. van Vollenhoven RF, Cifaldi MA, Ray S, et al. Improvement in work place and household productivity for patients with early rheumatoid arthritis treated with adalimumab plus methotrexate: work outcomes and their correlations with clinical and

Ref-9


167. Busija L, Pausenberger E, Haines TP, et al. Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQuOL). Arthritis Care Res (Hoboken). 2011 Nov;63 Suppl 11:S383-412. doi: 10.1002acr.20541. PMID: 22588759.


201. van Gestel AM, Anderson JJ, van Riel PL, et al. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American College of
