



Comparative Effectiveness Review Disposition of Comments Report

Research Review Title:

Drug Therapy for Early Rheumatoid Arthritis: A Systematic Review Update

Draft review available for public comment from December 20, 2018 to February 10, 2018.

Research Review Citation: Donahue KE, Gartlehner G, Schulman ER, Jonas B, Coker-Schwimmer E, Patel SV, Weber RP, Lohr KN, Bann C, Viswanathan M. Drug Therapy for Early Rheumatoid Arthritis: A Systematic Review Update. Comparative Effectiveness Review No. 211. (Prepared by the RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2015-00011-I for AHRQ and PCORI.) AHRQ Publication No. 18-EHC015-EF. PCORI Publication No. 2018-SR-02. Rockville, MD: Agency for Healthcare Research and Quality; July 2018. Posted final reports are located on the [Effective Healthcare Program search page](https://doi.org/10.23970/AHRQEPCCER211). <https://doi.org/10.23970/AHRQEPCCER211>.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site or AHRQ Web site in draft form for public comment for a 45-day period. This review was posted for 52 days with a 1-week holiday-related extension. Comments can be submitted via the Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.



Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Abstract	Abstract, line 36: "improved disease activity" – not exactly improved disease activity, improved disease control?	We agree and have made the change to 'improved disease control'.
Peer Reviewer #1	Evidence Summary	ES-3, lines 26-32 Clarify all DAS28 scores described are baseline mean or median scores of included studies?	The ranges of DAS28 scores described for each drug category capture both mean and median values of our included studies and we have clarified the text in this line.
Peer Reviewer #4	Evidence Summary	Page 14, Table ES-2, lines 10-14: unclear how the response is defined? Several response outcomes are mentioned, what if one changes and others didn't.	Response is defined by the ACR or DAS, and we have added footnotes to Table ES-3. If one changes and others do not, this is described in the detailed results. Additionally, Table 3 in the Methods describes the hierarchy of preferred measures for data abstraction.
Peer Reviewer #4	Evidence Summary	Page 13, lines 15-18: There is a wide range of prior DMARD and glucocorticoid use. There weren't any subgroup analysis that compared whether the results differed between different populations.	Although there is a wide range of prior DMARD and glucocorticoid use, we did not have a sufficient number of trials within each population comparison to do any subgroup analyses.
Peer Reviewer #4	Evidence Summary	Page 14, lines 45-48: Why was no strength of evidence provided for this comparison? Please consider adding.	Thank you for catching, we have added the strength of evidence.
Peer Reviewer #4	Evidence Summary	Page 16, lines 40-42: why speculate this as "possible"? Consider performing the Analysis of withdrawals due to inefficacy and compare between arms to substantiate/qualify this statement.	We reviewed the trials examining withdrawals due to lack of efficacy, and upon review with our statistician, we could not perform an analysis as there were only 1-2 trials per drug comparison that examined this. Thus, we can only describe these results qualitatively and did not perform quantitative analyses.
Peer Reviewer #4	Evidence Summary	Page 17, lines 49-51: how do you balance the higher efficacy with a burden of multiple drugs and potentially higher risk, with this as the first line therapy?	That is a good point, and we have added this point to the end of the paragraph under 'Findings in Relationship to what is Already Known'.
Peer Reviewer #4	Evidence Summary	Page 17, line 40-42: You are comparing a systematic review to a guideline/recommendations that incorporate systematic review, meta-analysis, patient values/preferences and experience of a panel of providers and patients with few hundred years of experience, where clinical trials do not exist. Is that a fair comparison? Might you be better off comparing to other contemporary systematic reviews instead? If you decide to still compare this review to ACR or EULAR recommendations/guideline, then perhaps a table showing	We agree our evidence is not directly comparable and have added this to the text in the ES. Under 'Discussion and Findings in Context' in the Evidence Summary, we note: "Our key findings may differ somewhat from the ACR guidelines for early RA for other reasons. This report assessed the comparative effectiveness based on current evidence. While not directly comparable, the ACR clinical guidelines moves beyond evidence to make recommendations when evidence is limited."

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		key similarities and differences from your review would help. My suggestion is that you compare apples to apples (systematic reviews) rather than apples to oranges (guideline).	
Peer Reviewer #4	Evidence Summary	Page 17, line 46-48: To be specific, this is DMARD monotherapy (MTX preferred) in the ACR guideline, not MTX monotherapy.	Thank you for catching this; we have revised the text.
Peer Reviewer #4	Evidence Summary	Page 18, line 3-12: I did not see the evidence for poor prognostic factors in your analyses. Which trials provide analyses of these factors? please present those data. An active disease as in the trails analyzed doesn't equate poor prognostic factors, as specified in the EULAR guideline. If trial data were analyzed or presented by these prognostic factors, present that analyses. If no such analyses exist, please remove this interpretation that your analyses showed support for this EULAR recommendation "specifically". From where I can see, this interpretation goes much beyond the data presented. does not support this.	We did not do formal analyses on studies with poor prognostic factors, but we described them when studies noted their patient population had poor prognostic factors (studies cited in that paragraph). The studies cited do support our statement.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Evidence Summary	Bristol-Myers Squibb (BMS) is a global biopharmaceutical company whose mission is firmly focused on discovering, developing, and delivering innovative, transformational medicines for patients with serious diseases. BMS has a deep and long standing commitment to immunotherapy that began over 20 years ago, and we continue to pioneer novel approaches to optimize the body's immune response. Our goal in autoimmune disease is to deliver life-changing medicines for patients. BMS is pursuing a wealth of Immunoscience research and is at the forefront of addressing unmet patient needs where treatment options are limited or improvements are needed. BMS is dedicated to advancing the science of immunology and to disseminating the results of our research to ensure that our work can benefit the widest range of patients.	Noted.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>On behalf of Bristol-Myers Squibb, please find our response to the Agency for Healthcare Research and Quality's call for public comment on the draft systematic review titled, <i>Drug Therapy for Early Rheumatoid Arthritis in Adults: A Systematic Review Update</i>. As requested, we have submitted our response electronically via the portal located on AHRQ's Website.</p> <p>Please note that Bristol-Myers Squibb does not recommend the use of a product in any manner inconsistent with that described in the ORENCIA full Prescribing Information (also attached).</p>	
<p>Public Commenter #4 (Leticia Ferri on behalf of BMS)</p>	<p>Evidence Summary</p>	<p>Page ES-2, Results and key findings</p> <p><u>Comment:</u> Authors state a total of 14 studies were reported as having high risk of bias and excluded from the main analysis. Additionally, the authors claim a sensitivity analysis was performed and reported in Appendix H accounting for these studies while Appendix H states only 2 studies with high risk of bias for sensitivity analysis. The information in this section is discordant and does not correspond to Appendix H.</p>	<p>Actually, 16 studies were reported as high risk of bias, but not all of these studies were eligible to be used in the NWMA. The sensitivity analyses were performed with 2 studies that were rated high ROB and also eligible for the NWMA.</p>
<p>Public Commenter #5 (Jason Spangler on behalf of Amgen)</p>	<p>Evidence Summary</p>	<p><u>Comment:</u> The evidence reported on drug effects on functional capacity does not include the totality of the available evidence.</p> <p>It should be noted that the COMET (COMbination of Methotrexate and ETanercept) trial² demonstrated that the combination of ETN plus MTX produced higher ACR response, higher remission, better radiographic outcomes, and greater functional capacity than MTX alone. This data fulfills the evidence requirements and should be included in the Report.</p> <p>If not included, then the rationale for specifically including only results of adalimumab and tocilizumab, but excluding evidence for other drug therapies should be provided.</p>	<p>We agree, and COMET is included in the final report and tables.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
<p>Public Commenter #5 (Jason Spangler on behalf of Amgen)</p>	<p>Evidence Summary</p>	<p>Comment: The Draft Report states “the combinations of either a TNF or a non-TNF biologic plus MTX also produced greater functional capacity, except for etanercept (a TNF) or abatacept (non-TNF), for which results were inconclusive.” We suggest the use of the word “except” is inaccurate and should be reworded to accurately express the evidence on ETN. The use of the word “except” suggests that ETN and abatacept have no evidence of producing greater functional capacity relative to controls or comparators. However,</p> <p>There is evidence for ETN regarding the assessment of functional capacity, and while limited, the results from the Enbrel Early RA (ERA) study³ show benefits or lack of difference of ETN depending on the timepoint reported, and as such are inconclusive in this early RA population.</p> <p>The statement should be reworded to accurately express the evidence on ETN. We suggest revising to read - “The combinations of either a TNF or a non-TNF biologic plus MTX also produced greater functional capacity. Results for ETN (a TNF) and abatacept (a non-TNF) were inconclusive.”</p> <p>A similar update should be made to the relevant Key Point on page 46 which suggests that combinations of other TNF biologics plus MTX do not produce statistically significantly greater improvements in functional capacity in comparison to MTX alone.</p>	<p>We re-reviewed the data and modified the sentence to: ‘The combinations of several TNF (adalimumab, certolizumab pegol, infliximab) and non-TNF biologics (rituximab) plus MTX also produced greater functional capacity. The results for the remainder of the biologics (etanercept, abatacept, tocilizumab) were inconclusive.’</p>
<p>Public Commenter #6 (Fang Sun on behalf of Merck)</p>	<p>Evidence Summary</p>	<p><u>ES-1:</u> As the authors pointed out, compared to the 2012 report, this report has a different scope of the disease, focusing only on early RA. Therefore, this report should not be viewed as an update to the 2012 report. As we comment below, due to the lack of agreement on the definition of early RA, existing evidence is suited better for addressing comparative effectiveness questions for the general RA population than for a poorly-defined early RA population. Therefore, we suggest PCORI adhere to the original scope of the 2012 report for this update, which would yield richer, less-biased information to guide RA management.</p>	<p>Although there is no consensus on the definition of early RA, the scope of the report was determined by stakeholder and expert input.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #6 (Fang Sun on behalf of Merck)	Evidence Summary	<u>Table ES-1</u> : Since patient adherence/discontinuation could be due to non-clinical reasons, total discontinuation may not be a good measure of harm.	We used overall discontinuations as this was used in the earlier report. We agree there are limits to this measure and thus additionally examined measures of discontinuation due to adverse events.
Public Commenter #6 (Fang Sun on behalf of Merck)	Evidence Summary	Although the analyses didn't support the combo therapy of MTX and biologic in all early RA, it did conclude the benefits of the combo therapy for early RA patients with moderate to high disease activities. This important finding should be appropriately noted in the report.	We agree and have modified the text in the report.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	Table ES-1: Population Included in the Review <ul style="list-style-type: none"> Inclusion Criteria: Key inclusion criteria for patients with RA is typically 18 and over as indicated by the definitions in the ACR guidelines. Additionally, patients enrolled in Actemra and Rituxan clinical trials were adults 18 years of age or older. 	We modeled our current search on our prior report searches. We captured patients aged 18 years and older in our searches and have updated our inclusion criteria in the report to clarify this.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	Table ES-1: Population Included in the Review [continued] <ul style="list-style-type: none"> Exclusion Criteria: Consider including patients for up to 2 years as also noted in the definition of early RA from Actemra and Rituxan clinical trials. 	There is no consensus on the optimal definition of early RA. We have added more justification to the reason we chose a shorter disease duration in the Introduction.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	Table ES-1: Key Questions Covered by the Review <ul style="list-style-type: none"> Disease activity and remission: Consider evaluating ACR as continuous measure of clinical response beyond ACR50, as well as including CDAI. 	This report included the most common outcome measures for RA. We were limited by the measures chosen by the individual studies. We chose not to add ACR as a continuous measure of clinical response and CDAI, as these were less commonly reported. If this report is updated in the future, we will consider these measures as their frequency of reporting increases.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	Table ES-1: Key Questions Covered by the Review [continued] <ul style="list-style-type: none"> Radiographic measures on slowing or limiting the progression of joint damage: Consider that radiographic measures differ by scoring methods and are not consistent across all clinical trials. There are various methods (ex, Sharp-van der Heijde method vs. Sharp/Genant vs. Larsen) 	The 2 measures are highly correlated (r-squared ranging from 0.95 to 0.99): (http://www.openaccessjournals.com/articles/comparison-of-the-genantmodified-sharp-and-van-der-heijdemodified-sharp-scoring-methods-for-radiographic-assessment-in-r.pdf). ¹ To account for the difference scales, we are using SMDs as our effect size comparing different drugs.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		used to define one drug's radiographic findings and making comparisons/extrapolations is not valid between two different drugs with different x-ray scoring systems.	
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	<p>Table ES-1: Key Questions Covered by the Review [continued]</p> <ul style="list-style-type: none"> • Patient-Reported Symptoms: Consider expanding PRO's to include additional outcome measures such as FACIT-F EQ-5D, PGA, MCS, PCS and IPQ-R. 	Agree – we did include the PROs outlined. Selected PROs were discussed in the text, and the remainder are found in the tables.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	<p>Table ES-3: Benefits and harms of biologic DMARDs for early RA treatment- Disease Activity</p> <ul style="list-style-type: none"> • <u>Response</u>: Consider modifying Actemra from insufficient to moderate strength of evidence based on significant improvement in response scores for Actemra combination and monotherapy vs. MTX. ACR was considered a secondary endpoint and there were significantly greater response rates observed for Actemra vs. MTX. Although some measures did not reach statistical significance for monotherapy in the FUNCTION trial, the ACR response rates still indicated improvement in RA signs and symptoms for monotherapy at weeks 24 and 52. In the U-ACT-Early trial, Actemra combination and monotherapy had a high ACR50 response as compared to MTX (statistical significance was not reported between combination and monotherapy). The proportion of patients with EULAR good response at week 24 was significantly greater in both Actemra combination and monotherapy vs. MTX. A similar response was seen for ACR(20/50/70/90) at weeks 24 and 52. 	After consideration, we have kept the overall strength of evidence as insufficient for the disease activity comparison due to inconsistent results and lack of precision.
Public Commenter #7 (Tom Innal on behalf of Genentech)	Evidence Summary	<p>Table ES-3: Benefits and harms of biologic DMARDs for early RA treatment- Disease Activity [continued]</p> <ul style="list-style-type: none"> • <u>Remission</u>: Consider changing Actemra from low to moderate strength of evidence for remission. Also consider modifying the statement to reflect the data in Actemra monotherapy as there is minimal difference between Actemra combination therapy and monotherapy in regards to clinical remission. In both clinical trials, a significantly greater 	<ul style="list-style-type: none"> • We have reviewed this evidence for the comparison (Significantly higher remission for non-TNF biologic (TCZ) + MTX than TCZ or MTX alone^{2, 3}) and downgraded the strength of evidence from moderate to low due to precision (large confidence intervals that cross appreciable differences or harms). We have added the 24-week primary outcomes to the Results section as suggested.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>proportion of patients in the Actemra combination and monotherapy arms achieved DAS28-ESR remission as compared to MTX ($p < 0.0001$). For the U-ACT-Early trial, we recommend focusing on the results with the initial treatment regimen. Sustained remission was not different between the combination vs. the monotherapy arm ($p = 0.62$). The p value of 0.06 which was used for the analysis in Table 6 is incorrect since this is for the three treatment strategies during the entire course of the study (ex. initial plus subsequent other drug regimens for 104 weeks) and should be changed to $p < 0.0001$ to reflect the initial treatment regimen.</p>	<ul style="list-style-type: none"> In the Results section, Non-TNF Biologic: MTX Plus Non-TNF with Either MTX or Non-TNF Biologic, we added the following: '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%, $p < 0.001$)' We also added this to what is now Table 7.
<p>Public Commenter #7 (Tom Innal on behalf of Genentech)</p>	<p>Evidence Summary</p>	<p>Table ES-3: Benefits and harms of biologic DMARDs for early RA treatment- Disease Activity [continued]</p> <ul style="list-style-type: none"> Radiographic Progression: Modify the statement for Actemra to reflect data supporting that Actemra monotherapy also produced less radiographic progression as compared to MTX alone. Radiographic joint damage progression was low in all treatment arms, but at week 104 it was significantly less in both Actemra arms compared with the MTX arm. Combination therapy with Actemra produced slightly higher remission and less radiographic progression as compared with monotherapy but this was not statistically significant between the two arms and they should be considered comparable. Recent data from the U-ACT-Early trial evaluated the effect of study treatments on erosive joint damage and joint space narrowing. The mean changes from baseline in Sharp-van der Heijde score (used to evaluate radiographic progression) were significantly lower for the Actemra combination group than for the MTX group at week 52 ($p = 0.016$) and at week 104 ($p = 0.021$), and they were significantly lower for the Actemra monotherapy group than for the MTX group at week 104 ($p = 0.038$), but not at week 52 ($p = NS$). Mean changes from baseline to week 104 in total erosion scores were significantly lower in both the Actemra combination ($p = 0.016$) and monotherapy group ($p = 0.023$) than in the MTX group. Mean changes from baseline in total joint-space narrowing scores did not differ significantly between treatment groups at weeks 52 or 104. At week 104, 	<p>Thank you for the information. We have added the 2 year radiographic findings to Table ES-3.</p> <ul style="list-style-type: none"> Significantly less radiographic progression for non-TNF biologic (TCZ) + MTX than MTX (but not TCZ) alone^{2,3}

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>there were also significantly fewer erosions of the feet in the Actemra combination (p=0.046) and monotherapy group (p=0.022) than in the MTX group. Note that in the FUNCTION trial, the Actemra arms were only compared to MTX, not each other, and statistical significance was not measured as there was not enough power to detect differences between the two Actemra arms. Additionally, Actemra combination therapy with MTX included both 8 mg/kg and 4 mg/kg and these doses produced different results. The findings from the study describing less radiographic progression in both Actemra arms with statistical significance was reached with combination therapy vs. MTX due to the hierarchical chain break.</p>	
<p>Public Commenter #7 (Tom Innal on behalf of Genentech)</p>	<p>Evidence Summary</p>	<p>Table ES-3: Benefits and harms of biologic DMARDs for early RA treatment- Functional Capacity</p> <ul style="list-style-type: none"> • <u>Functional Capacity</u>: Consider modifying Actemra from insufficient to moderate strength of evidence for functional capacity as there is recent data to support the improvement of PROs with Actemra combination or monotherapy vs. MTX. QOL outcomes (including SF-36, EQ-5D, FACIT-F, and IPQ-R) from the U-ACT-EARLY trial were recently published. Compared with the mean score for the MTX group, significantly greater improvements over time in mean SF-36 PCS were reported in the Actemra monotherapy (p=0.012) and in the Actemra combination group (p=0.044). In contrast, no statistically significant differences over time were reported between treatment groups for SF-36 MCS. At weeks 12 and 52, the proportion of patients who achieved MCID in SF-36 PCS was 75.6% (p=0.016 vs. MTX) and 88.5% (p=0.03 vs. MTX), respectively, in the Actemra monotherapy group; 72.7% (p=0.049 vs MTX) and 88.7% (p=0.027 vs. MTX), respectively, in the Actemra combination group; and 58.6% and 72.9%, respectively, in the MTX group. The proportion of patients who achieved MCID in SF-36 MCS was not significantly different between groups (p≥0.06). Additionally, significantly greater improvements over time in mean EQ-5D scores were reported in the Actemra combination group than 	<p>We appreciate the feedback. We have added the new PRO data for the U-Act-Early study into the report. After further consideration, we have kept the strength of evidence insufficient. The overall evidence for this comparison was mixed for functional capacity (ranging from no significant differences in functional capacity to significant differences in functional capacity).</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>in the MTX group (p=0.018). At Week 24, the proportion of patients who achieved MCID in EQ-5D was 72.8% in the Actemra combination group (p=0.045 vs. MTX), 69.1% in the Actemra group (p=NS vs. MTX), and 58.6% in the MTX group. Except for the identity domain in the IPQ-R (Actemra vs. MTX, p=0.048), no other statistically significant between-group differences over time were noted for IPQ-R or FACIT-F. PRO data from the FUNCTION trial demonstrated statistically significantly greater improvement in HAQ-DI scores from baseline to weeks 24 and 52 for Actemra combination therapy vs. MTX (p=0.0011 and 0.0024, respectively). Percentages of patients with the MCID of ≥ 0.22 in HAQ-DI at week 52 were numerically higher for Actemra combination therapy than for MTX. Mean improvements in FACIT-F, SF-36 PCS, SF-36 MCS, Pain VAS, and PGA VAS scores at weeks 24 and 52 were larger for Actemra combination vs MTX. Clinically relevant improvements of ≥ 5 for FACIT-F, >5.42 for SF-36 PCS, and >6.33 for SF-36 MCS were reached by week 24 with Actemra combination therapy and were maintained through week 52. Numerically higher improvements were also observed across all endpoints for Actemra combination therapy vs. MTX. All PRO responses for Actemra monotherapy were largely similar to those seen for MTX.</p>	
Peer Reviewer #1	Introduction	Page 4, lines 26-32 The logic of the first part of the paragraph vs the last sentence does not follow?	We have modified the text to incorporate the same subject (drugs) into the first and last sentence.
Peer Reviewer #2	Introduction	As above, justify scientifically or conceptually (or was it just practical due to the volume of data) why you chose to split out early disease.	We have added more description surround the context of the definition in the introduction. 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. Given the above caveats and limitations of placing boundaries on the continuum of early RA, this is the basic definition (no more than 1 year of diagnosed RA)

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
			we adopted for this systematic review update. The goal of separating early from late disease, however one defines these stages, is not to assess whether certain therapeutics might be more effective in early versus late disease, but to provide some rationale to physicians and patients regarding an evidence-based approach in early disease.
Peer Reviewer #2	Introduction	Include a table of definitions so the reader doesn't have to hunt to find where a particular one was initially defined.	We have added a table defining all of the report's abbreviations and acronyms. It appears immediately after the main report/conclusion section.
Peer Reviewer #2	Introduction	Overall the Introduction is good, and the key questions identified are indeed important ones.	Thank you.
Peer Reviewer #2	Introduction	On page 1 under Introduction, the potency of the MHC Class II 'shared epitope' should be called out as the driving genetic force in RA, rather than implying that 100 identified polymorphisms have equivalent potency.	We have modified the text to reflect this.
Peer Reviewer #2	Introduction	On page 3, in paragraph beginning with 'The optimal initiation strategy.....', the implication is that only 1 of the 3 strategies can be chosen. However, it is well accepted that the overarching principle should be 'treat-to-target'. How one gets there is more debatable - e.g., step-up vs step-down. In this paragraph, the implication that they are mutually exclusive choices should be modified.	We have modified the text to reflect that the overarching principle should be to 'treat to target'.
Peer Reviewer #3	Introduction	I was pleased to see that early in the report patients are clearly identified as healthcare decision-makers in addition to clinicians and others. (page ii, line 17)	Thank you.
Peer Reviewer #3	Introduction	Nulliparity was described as a risk factor for RA (page 1, line 20) and I'm just curious about whether that has been fully evaluated? Is it the case that people with RA--for a variety of reasons--end up not having their own children, or is it actually thought to be causal?	It is thought to be causal, but the evidence is based on old epidemiologic studies. ⁴ Thus, it is not really clear. In the text, we note that the etiology is incompletely understood and have decided to leave it at that, rather than speculate.
Peer Reviewer #4	Introduction	This seems reasonable.	Thank you.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
<p>Public Commenter #4 (Leticia Ferri on behalf of BMS)</p>	<p>Introduction</p>	<p>Page 1, Definitions of Early RA and Challenges with the Definitions</p> <p><u>Recommendation:</u> Please consider extending the definition of early RA from ≤ 1 year to ≤ 2 years from diagnosis.</p> <p><u>Rationale:</u> The definition of early RA has evolved over the years. Most recently, early RA has been defined as < 6 months from onset of symptoms [1] and within 3 months of onset of symptoms [2]. Previous to these definitions, early RA was defined as ≤ 2 years of disease duration (from diagnosis). This definition is still used by the National Institute for Health and Care Excellence (NICE) guideline on the management and treatment of RA (updated most recently in 2015) [3]. Thus, some randomized clinical trials in early RA included patients of duration less than two years, since this was a commonly accepted definition at the time of trial design.</p> <p>We recognize the lack of consensus on the onset and duration of early RA and agree with the European taskforce in that despite ACR's definition of early RA as < 6 months, for clinical research purposes (patient recruitment and population characterization) this duration of time from diagnosis needs to be extended. Others have stated, "It has been postulated that the early stages of RA may, therefore, offer a therapeutic window of opportunity in which to prevent joint damage from occurring; it has been suggested that this window may exist from 3 months to 2 years following the onset of symptoms." [4]</p> <p>Furthermore, we note that the AHRQ-PCORI systematic review states "Most of our trials of biologics (n=10) enrolled mixed populations of early RA patients and those with longer-duration RA." [5] However, 11 papers were excluded on the basis of being eligible, except the definition of early RA was up to 2 years. Inspecting these papers further reveals that at least 5 trials clearly include mixed populations with the majority of participants having a duration of ≤ 2 years. Given the relatively small number of trials included in the network,</p>	<p>After careful consideration, we opted to keep the definition of early RA to 1 year of disease. We have added additional justification for this in the Introduction.</p> <p>All included populations were of early RA less than 1 year. We included mixed populations if more than 50% of the study population had an early RA diagnosis.</p> <p>We scanned the 11 studies and found that 3 studies were more appropriately excluded for mean duration of RA over 2 yrs, and 1 study for being an abstract-only record. To reflect this change in the number of these studies, we have updated the count presented in the text to 7 studies (reported in 10 articles). In these 7 studies, we did not find any differences from our current report's eligible evidence base. We added a brief summary statement to the Evidence Summary (Limitations to the Evidence Base - second to last paragraph) and Main Report (limitations section of the Discussion - second paragraph).</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>adding these trials may improve precision of the estimates, without unduly affecting validity.</p> <p>For these reasons, we propose that AHRQ-PCORI's SR further extend the definition and include trials that defined early RA as ≤ 2 years disease duration since diagnosis. In addition, AHRQPCORI should also consider RCTs that do not have disease duration as an inclusion criterion but have disproportionately recruited patients with ≤ 2 years of disease duration (i.e. $> 50\%$ of the population, which we also note was arbitrarily chosen with no strong rationale).</p> <p>Expanding the definition of early RA to include patients ≤ 2 years from diagnosis might also increase the proportion of patients with early RA and other subgroup population characteristics (e.g., poor prognostic factors or disease activity), thereby permitting further subgroup analyses.</p>	
<p>Public Commenter #4 (Leticia Ferri on behalf of BMS)</p>	<p>Introduction</p>	<p>Page 7 and 8, Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions</p> <p><u>Recommendation:</u> Search EULAR and ACR conference websites for the period of 2014 to 2017 for unpublished (or ongoing) trials with results.</p> <p><u>Rationale:</u> EULAR and ACR usually contain abstracts and posters on key yet unpublished trials. As part of increasing the sensitivity of the searches and covering grey literature, it is a common practice, for conducting major systematic literature reviews, to search for main conferences and proceedings published during the past 2-3 years. The current systematic literature review did not perform this search. This becomes important especially in the context of defining early RA as per recent claimed recommendations since it takes a few months to a few years before a major trial (with certain definition of RA for instance) gets published in a peer-reviewed journal and then catalogued by MEDLINE or PubMed and we believe that by covering important conferences such as EULAR and ACR, the authors of this report could have captured populations which were more a match in terms of the set</p>	<p>We search the gray literature, which also included abstracts. Abstracts were excluded given high risk of bias and/or limited information.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		definition of early RA and thus could have been better able to conduct analysis based on more recent available evidence.	
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Introduction	<p><u>References:</u></p> <p>[1] Singh JA, Saag KG, Bridges SL, Jr., et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. <i>Arthritis & rheumatology</i> (Hoboken, NJ). 2016;68(1):1-26.</p> <p>[2] Smolen JS, Landewe R, Bijlsma J, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. <i>Ann Rheum Dis</i>. 2017;76(6):960-977.</p> <p>[3] National Institute for Health and Care Excellence. Rheumatoid arthritis in adults: management. [Practice Guideline]. 2015; https://www.nice.org.uk/guidance/cg79/evidence. Accessed 2/2/2018.</p> <p>[4] Fleischmann RM, Huizinga TW, Kavanaugh AF, et al. Efficacy of tofacitinib monotherapy in methotrexate-naive patients with early or established rheumatoid arthritis. <i>RMD open</i>. 2016;2(2):e000262.</p> <p>[5] Agency for Healthcare Research and Quality P-CORI. Drug Therapy for Early Rheumatoid Arthritis: A Systematic Review Update. 2017; https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/draft-report-drug-therapy-for-earlyrheumatoid-arthritis.pdf. Accessed 2/2/2018.</p>	<p>Noted. We used most of these references, and one referred to our ongoing report.</p> <ol style="list-style-type: none"> 1) <u>Singh, 2016</u>: BKG reference 2) <u>Smolen, 2017</u>: BKG reference 3) <u>NICE, 2015</u>: We originally cited the 2009 version of these guidelines, but we have changed it to the 2016 guidelines in the report instead. 4) <u>Fleischmann, 2016</u>: This is a post-hoc analysis analyzing the subgroup of patients with early RA (<1 year duration). It uses data from the ORAL Start trial of TOF vs. MTX, which we excluded previously because ≥50% patients have RA >2 yrs duration based on the sample's average disease duration. 5) <u>AHRQ/PCORI, 2017</u>: Our ongoing report
Public Commenter #6 (Fang Sun on behalf of Merck)	Introduction	P1: As the authors acknowledged, no consensus exists on the onset and duration of early RA. For this review, the authors defined early RA as no more than 1 year of diagnosed disease duration. But a quick review of the study population information provided in Table C-1 suggests that more than half of the studies included in this review (25 out of 46) enrolled both early and post-early RA patients. Below is a summary of the patient enrollment criteria used in the 25	In the earlier report, early RA was defined as less than 3 yrs. Based on current literature and rheumatology experts, we chose one year or less and allowed populations in which >50% of patients had early RA. As more evidence becomes available, we will consider making the definition more strict (e.g. >75% with early RA 1 year or less) if this review is updated in the future.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>studies in terms of diagnosed disease duration (in parenthesis)</p> <ol style="list-style-type: none"> 1. The Enbrel ERA study (<3 years) 2. The PROWD study (<2 years) 3. Bili et al., 2014 (not specified) 4. Bliddal et al., 2015 (not specified) 5. The COBRA study (< 2 years) 6. The PREMIER study (<3 years) 7. The FUNCTION study (< 2 years) 8. the CARDERA study (< 2 years) 9. Conaghan et al., 2016 (≤2 years) 10. Cummins et al., 2015 (<2 years) 11. The COBRA-light study (<2 years) 12. The COMET study (3-24 months) 13. The AVERT study (<2 years) 14. The BeST study (<2 years) 15. The IMPROVED study (≤ 2 years) 16. Lan et al., 2017 (not specified) 17. The NEO-RACo study (not specified) 18. The ERAN study (not specified) 19. The TEAR study (< 3 years) 20. The FIN-RACo study (<2 years) 21. The ORBIT study (not specified) 22. The ASPIRE(≤3 years) 23. The IMAGE study (8 weeks to 4 years) 24. The HOPEFUL 1 study (≤ 2 years) 	

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>25. The AGREE study (≤ 2 years)</p> <p>It appeared that the authors included in this review all studies with $>50\%$ early RA that also met the other inclusion criteria. However, 50% is a very low threshold for selecting early RA studies. Applying a higher threshold (e.g., 85% or 90%) is essential to ensure the findings of the review generalizable to the early RA population.</p>	
Peer Reviewer #1	Methods	<p>Description of Data synthesis and network MA is clear. However, in NWMA, at most only 2 studies were included in each comparison and in most comparisons, there is only one study (while the pairwise MA needs a minimum of 3 studies). While the authors showed great efforts to ensure the transitivity assumption, the included studies don't provide adequate data to check on the consistency of NW.</p> <p>Within FDA approved dose ranges, are there substantial dose variations across studies? Are there dose-response relationship within the dose range? What is the dose range of MTX in studies included in NWMA?</p> <p>Please provide the length of follow up for studies included in NWMA.</p>	<p>We agree, the dearth of data is a limitation of our analyses. In cases of NWMA without closed loops, no statistical assessment of consistency can be made. We address this limitation in the Discussion.</p> <p>Within FDA-approved ranges, there were no substantial variations in average dosages. Biologics have standardized dosing regimens per kg body weight.</p> <p>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.</p>
Peer Reviewer #1	Methods	<p>For pairwise meta-analysis, it would be fine to combine low or medium ROB studies with high ROB studies, and investigate the impact of high ROB studies in a sensitivity analysis. In many cases, high ROB studies have not led to much differences in the results.</p>	<p>If we included high ROB studies for pairwise meta-analyses, we would have enough studies to conduct meta-analyses. However, we would not have enough low and medium risk of bias studies to conduct sensitivity analyses without high ROB studies. In other words, we would not be able to assess the impact of high ROB studies on the estimates of effect. Therefore, we decided against pairwise meta-analyses.</p>
Peer Reviewer #1	Methods	<p>Please provide information on which version of SMD was used in the analysis.</p>	<p>The SMD in the plots is the standardized mean difference (i.e., mean difference divided by standard deviation). We add this description to the Methods text under "Data Synthesis" and each radiographic forest plot.</p>
Peer Reviewer #1	Methods	<p>Page 6, inclusion criteria: Placebo controlled trials are only considered for NWMA?</p>	<p>Because our report is a comparative effectiveness review, we did not focus on the general efficacy of</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
			treatments. Placebo-controlled trials would have been eligible for the NWMA; however, we did not detect any placebo-controlled trials for our population of interest.
Peer Reviewer #1	Methods	Page 7 lines 29-31 Studies included mixed population with < 50% of early RA , but reported results separately of early RA were also included?	We included any study with mixed populations if analyses stratified results by early and established RA. If results were not stratified, we excluded studies in which less than 50% of participants had early RA.
Peer Reviewer #1	Methods	Table 3: All reported bolded outcome measures will be abstracted, if a study reported more than one?	We abstracted all bolded outcomes, even if a study reported more than one.
Peer Reviewer #1	Methods	Page 10 line 10: Clarify what items about the statistical analysis of RCTs were added? (For example, methods when baseline imbalance was observed? Or something else?)	We clarified in the text that we added an item about intention-to-treat analyses.
Peer Reviewer #2	Methods	The Methods are stringent and well thought out and justified. I cannot comment meaningfully on the statistical methods and will leave that to others, but I do like the addition of the network analyses. And the decisions regarding how to manage studies with low SOE and/or high ROB are thoughtful and defensible.	Thank you and noted.
Peer Reviewer #2	Methods	As above, I think it would have been much easier and relevant to have separated out treatment naïve as the early population, and treatment experienced as the 'late' disease population.	Our NWMA separates out the treatment naïve studies. In the results, we have tried to make this clearer and group treatment naïve and treatment resistant studies.
Peer Reviewer #2	Methods	Regarding K3, it would have been helpful to split out Serious Infections and Malignancies as adverse events of particular interest to rheumatologists and rheumatology patients (if these were identifiable in most publications). These drive a lot of decision making by both patients and doctors. Did sponsorship of the studies (pharma vs non-pharma) factor into the ROB assessment?	Data were sparse on comparative differences in serious infections and malignancies in this early RA population. We have added this to the Discussion under 'Applicability'. As noted under 'Assessment of Methodologic Risk', we used the criteria from the ROBINS-I and Cochrane ROB tools. Sponsorship of studies did not factor into the ROB assessment.
Peer Reviewer #3	Methods	I previously addressed the challenge of the definition of "early" RA and the potential that patients described this way may actually have longer-term disease. (multiple references)	Noted, and we added this as a limitation in the discussion.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Methods	I was puzzled by the placement of patient adherence as a sub-part of key question 3 (page 4, lines 50-53) which addresses "harms" versus key question 2 which addresses QOL, PROs, and functional capacity. It would seem to frame (lack of) adherence as a reaction to side effects (whether serious or not), rather than as part of the larger equation about how the medication protocol appropriate fits in to the patient's life (does the medication dosing "fit in" to the patient's schedule, are the side effects manageable enough that the patient tends to be adherent, are the patients' treatment goals and preferences addressed appropriately by the medication, etc.).	We viewed lack of adherence as an undesirable event. This is the reason why we present it in the harms chapter. We agree that lack of adherence can have other reasons than harms (e.g., lack of efficacy).
Peer Reviewer #3	Methods	Throughout the document, studies involving MTX are described and results provided, but it wasn't clear what doses were used. From a patient perspective, MTX can be a tough drug with escalating side effects according to dose. When evaluated in head to head comparison with other drugs, it would be really valuable for patients to have a sense for the doses used.	We report MTX doses in the in-text tables of the Results.
Peer Reviewer #3	Methods	I could not tell from the methods section whether patient partners were involved in helping to develop the report (set inclusion criteria, develop key questions, synthesize data, etc.)? Having that patient lens is always so valuable to any research and publication process. I was struck for example by the explanation of the exclusion of Anakinra (page 6, lines 33-34) and wondered if any patients might have provided insights on whether there are occasions where it is prescribed, even if off-label or under unusual circumstances.	Yes. Further details about our patient partners can be found in Appendix J. We also added some descriptive text to the Methods, second paragraph. Although anakinra is an FDA approved treatment for RA, it has not shown strong efficacy when compared with other DMARDs. Also of note, the 2015 ACR guidelines did not include anakinra because of its infrequent use in RA and lack of new data since 2012 (Singh et al., 2015). ⁵ We do not have any insights on off-label use of anakinra.
Peer Reviewer #4	Methods	Inclusion and exclusion criteria are reasonable. Search strategies are logical.	Thank you.
Public Commenter #4 (Leticia Ferri on	Methods	Page 11, Data Synthesis: network meta-analyses using a multivariate random effects meta-regression model with restricted maximum likelihood estimation	The previous report employed a Bayesian framework with flat priors (i.e., non-informative priors), which is essentially equivalent to a frequentist NWMA. The reason to change the model was simply for convenience

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
behalf of BMS)		<p><u>Comment:</u> The original systematic literature review (SLR) and network meta-analysis (NMA) [1] used Bayesian framework via multiple treatment comparison meta-analysis using the methods developed by the Multi-Parameter Evidence Synthesis (MPES) Research Group at the University of Bristol for the analysis, however, it seems that the authors used a different package for the updated review (the current one) using a Frequentist approach without bringing a rationale as to why a different approach was used.</p> <p><u>Recommendation:</u> Use Bayesian framework for both random and fixed effect network meta-analysis. Consider using the deviance information criterion (DIC) to choose the model selected. Please describe clearly (with supporting rationale) the statistical method chosen for this analysis and if there was any adjustment for covariates in the model. Additionally, consider using a baseline risk meta-regression to assess the heterogeneity in the placebo response across studies.</p> <p><u>Rationale:</u> The earlier SLR used a Bayesian method for the NMA. There was no description provided for switching to frequentist method or justification for random or fixed effects models. NMA can be performed within a frequentist or Bayesian framework providing broadly comparable odds ratio for safety and efficacy. Bayesian methods involve a formal combination of a prior probability distribution (that reflects a prior belief of the possible values of the model parameters) with a likelihood distribution based on the observed data to obtain a posterior probability distribution of model parameters. The likelihood informs us about the extent to which different values for the parameter of interest are supported by the data. A major advantage of the Bayesian approach is that the method naturally leads into a decision framework [2-4]. The posterior distribution can be interpreted in terms of probabilities (e.g., “There is an x% probability that treatment A results in a greater response than treatment B”); frequentist approaches do not allow such an interpretation. By allowing calculation of</p>	<p>because of easier to use statistical software for frequentist NWMA.</p> <p>Results of Bayesian NWMA with flat priors and frequentist NWMA provide the same results.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>rank-probabilities, Bayesian NMA delivers a probability in terms of choosing best treatment option therefore making it a more flexible and a reliable method which offers more meaningful clinical interpretation of the results [5,6]. Therefore, for the sake of consistency with previous work as well as facilitation in clinical interpretation of the results we suggest using Bayesian over frequentist framework.</p> <p><u>References:</u></p> <p>[1] Donahue KE, Jonas DE, Hansen RA. Drug Therapy for Rheumatoid Arthritis in Adults: An Update. Agency for Healthcare Research and Quality (US); 2012 Apr. Report No.: 12-EHC025-EF</p> <p>[2] Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. <i>Statistical methods in medical research.</i> 2001;10(4):277-303.</p> <p>[3] Luce BR, Claxton K. Redefining the analytical approach to pharmacoeconomics. <i>Health economics.</i> 1999;8(3):187-189.</p> <p>[4] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health care evaluation. Chichester ; Hoboken, NJ: Wiley; 2004.</p> <p>[5] Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple treatment meta-analysis: an overview and tutorial. <i>Journal of clinical epidemiology.</i> 2011;64(2):163-171.</p> <p>[6] Carlin BP, Hong H, Shamliyan TA. Case Study Comparing Bayesian and Frequentist Approaches for Multiple Treatment Comparisons. Agency for Healthcare Research and Quality (US); 2013 Mar. Report No.: 12(13)-EHC103-EF.</p>	
Public Commenter #6 (Fang Sun on behalf of Merck)	Methods	<p><u>P6 Table 2:</u> If studies with >50% early RA are included in the analyses, this review may include a large proportion of post early stage RA patients, which would significantly affect the generalizability of the review's findings. See our previous comment.</p>	We have added this to the limitations.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #6 (Fang Sun on behalf of Merck)	Methods	<u>P9 Table 3</u> : Overall risk of harm is not explicitly defined in the table. Please provide a definition.	Based on the CONSORT statement, we define harms as “the totality of possible adverse consequences of an intervention or therapy; they are the direct opposite of benefits, against which they must be compared.” ⁶
Peer Reviewer #1	Results	Describe the implications of the network structure on the validity of the network meta-analysis.: The network plot is presented, and no discussion on the structure. The network structure is mainly star like, indicating there is no "real" network and the included studies don't provide adequate data to check on the consistency of NW.	We added text describing the network structure and its implications in the Results whenever we present network diagrams. We also added more text to the 'Limitations' section in the Discussion.
Peer Reviewer #1	Results	Describe the implications of the network geometry on the validity of the network meta-analysis.: No summary of the network geometry. The major concern is that only one or two studies were included in each comparison for both KQ1 and KQ3. There were not even enough studies to evaluate heterogeneity among studies. As the investigators also acknowledged in the limitation of the discussion section, most networks are star like, there is very limited information to evaluate the consistency of network.	We added text describing the network structure and its implications in the Results whenever we present network diagrams. We also added more text to the 'Limitations' section in the Discussion.
Peer Reviewer #1	Results	Describe the implications of the result of the analysis (or of diagnostics) on the validity of the network meta-analysis.: Appendix G provided results on test of consistency. Based on network plots, there are three closed loops and it is not clear why the test of consistency is only based on one loop or done correctly?	We have revised the consistency tables to include the results of direct and indirect comparisons for the closed loops in the network. Please note that only two of the three closed loops could be used in these comparisons. The final loop permits only direct comparisons because the treatments comprising that loop were assessed in the same two trials.
Peer Reviewer #1	Results	Describe the implications of the synthesis of results on the validity of the network meta-analysis.: Please see other comments.	The lack of head-to-head data and the general limitations of the NWMA are reflected in the strength of evidence (SOE) grades, which are mostly low or insufficient. In a few instances, the grade of SOE is moderate. In each case, we have at least one large, well-conducted RCT, and results from the NWMA are consistent with results from the RCT. In other words, in these cases, indirect comparisons of two interventions confirm the findings of direct comparisons.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	For KQ1 and KQ2, for each treatment/comparison, good descriptions were provided for the overall characteristics of included studies. However, given that the majority of included studies are RCTs, adding some summary comments on whether the patient characteristics were similar by randomized groups would be very informative to understand the included studies and evaluate the goodness of randomization.	Patient characteristics in most included RCTs were similar by randomized group. Studies in which characteristics between arms were different were rated as having a higher ROB.
Peer Reviewer #1	Results	Please provide more information on how the included studies met or unmet the inclusion criteria for NWMA. For Corticosteroids, there are typically multiple studies included (for example, 4 to 6 studies), but none of these studies is included in the NWMA. They were not even considered for a pairwise MA. Please clarify. (The limitation of the discussion section mentioned some heterogeneity issue, though not sure that is the main consideration, since conclusions were made on the level of drug category).	To ensure the transitivity assumption, studies had to meet the following criteria: (1) patients with early RA had not failed a prior treatment attempt 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. As for corticosteroid studies, we determined it was only feasible to include one study ⁷ found through our update literature searches in NWMA. The others were not eligible for NWMA or pairwise MA for a variety of reasons, including that several evaluated treatment strategies (rather than specific drugs), some were open-label RCTs, one lacked a clear comparator, and another was a single-arm study.
Peer Reviewer #1	Results	On the other hand, for NWMA, only one or two studies were included in each comparison for both KQ1 and KQ3. There were not even enough studies to evaluate heterogeneity among studies. Further, as the investigators also acknowledged in the limitation of the discussion section, most networks are star like, not “real” network, and there is very limited information to evaluate the consistency of network. When such limited information did not detect inconsistency, by no means it would be considered as evidence that the network is consistent, and the power to detect inconsistency is very low, too.	We agree, that our NWMA have limitations. Most of the results were derived from indirect comparisons relative to MTX, rather than mixed treatment comparisons. We acknowledge these limitations in the Discussion. The limitations are also reflected in the low or insufficient strength of evidence grades for these effect estimates.
Peer Reviewer #1	Results	Appendix G provided results on test of consistency. Based on network plots, there are three closed loops and it is not clear why the test of consistency is only based on one loop. In the text, the two reported coefficients are the same since they are	This was an oversight. For the final report, we conducted tests of consistency for all closed loops.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		from the same loop. Further, the direct and indirect point estimates were not close in Appendix Tables 2-4 and there is no power to detect any differences given the same number of studies.	
Peer Reviewer #1	Results	If the test of consistency is based on one loop only, there is barely no information to test the consistency of the overall network. In effect, the NWMA did many indirect comparisons based on single studies. In this case, the available data don't justify a formal NWMA and produce all pairwise comparisons as indicated in Figure 4, Figure 6 etc.	As noted above, we have added tests of consistency for the closed loops that permit both direct and indirect comparisons. We have added text noting the limitations of the NWMA given the available number of studies.
Peer Reviewer #1	Results	The similarity of event rates of each treatment across trials, if available, could be assessed to get a sense of the transitivity assumption. This is particularly true for the MTX arm since all included studies have a MTX arm. Important and selective indirect comparisons may be conducted when data justify.	We applied fairly strict criteria to ensure the transitivity assumption (which contributed to the dearth of data for our networks). These criteria are outlined in the Methods. Differences in event rates should be mitigated by the use of a relative outcome measure (relative risk).
Peer Reviewer #1	Results	Some of the conclusions based on NWMA may be stronger than it should be. For example, KQ1, "The TNF biologic ADA plus MTX had significantly higher ACR50 response (disease activity), smaller radiographic changes, and higher remission rates than ADA monotherapy (moderate SOE, supported by NWMA)." – however, such conclusion is based on 3 study loop with only one trial in each loop.	In cases where SOE grades are moderate, we have at least one large, well-conducted RCT, and results from NWMA are consistent with results from the RCT. In other words, in these cases, indirect comparisons of two interventions confirm the findings of direct comparisons. In the case of this specific example, plausibility supports our findings. Combination therapies of a biologic plus MTX are in general more efficacious than monotherapies. This is also the case in patients with established RA.
Peer Reviewer #1	Results	For the reporting of NWMA, the reported RR in the text (A vs. B) is often the 1/RR shown in the forest plot (B vs. A).	The use of the 1/RR was primarily for comparisons of combined therapies to MTX monotherapy. We have added plots of these comparisons for each outcome to make the text consistent with the graphical representations.
Peer Reviewer #1	Results	Page 34, lines 15-16, IFX plus MTX? Instead of ETN plus MTX?	Thank you for catching this. We have corrected the text to IFX plus MTX.
Peer Reviewer #1	Results	Page 35, lines 23-24 "The NWMA did not find any significant differences in ACR50 response for the combination of ABA plus MTX vs. ABA monotherapy (RR, 1.19; 95% CI, 0.95 to	After review of the NWMA plots, this was a mistake (ADA instead of ABA). This has been removed. The NWMA did find differences for ABA plus MTX vs. MTX and is

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		1.46)." - The direct evidence only comes from one trial and is this conclusion consistent with the results from trial (AVERT?)	consistent with the results from AGREE and AVERT. We have modified the text accordingly.
Peer Reviewer #1	Results	Page 36 section of Network meta-Analysis: please provide the number of studies included in NWMA here. Again there are not much data to evaluate the differences between the consistency and inconsistency models.	We added these numbers to the text.
Peer Reviewer #1	Results	There seems to be the potential to conduct pairwise MA to facilitate more reliable conclusions for the drug category in both KQ1 and KQ2, though it is hard to judge without additional information. For example, KQ1, Corticosteroids Versus csDMARDs, the direction of effects seem to be consistent for the included studies for disease activity and ACR response. In some comparisons of csDMARDs, and other drugs, there are also multiple studies.	We looked into pairwise MAs, but we almost never had data from at least three similar studies for any of the comparisons. One comparison, IFX plus MTX vs. MTX, did have 3 similar studies reporting the same outcomes that could be used in pairwise MA, but we determined it would not have been informative to synthesize them this way. The issue was that the analyses would have given excessive weight to one study whose sample size (1049) far exceeded the sample sizes of the other two (44 and 20, respectively).
Peer Reviewer #1	Results	In KQ3, there are statements like "Overall, no significant differences were found in discontinuations attributed to adverse events and serious adverse events." --- But there are five trials. Even if no trial showed no significant difference, could combining the five trials change the conclusion? Similar question applies to the section of csDMARD Combination Therapy Versus csDMARD Monotherapy, and maybe Adalimumab under TNF biologics (Page 69) in KQ3.	As noted above, we looked into pairwise MAs, but we did not have data from at least three similar studies for any of the comparisons.
Peer Reviewer #1	Results	Please provide more information to explain (why or why not pairwise MA in the specific situations).	As noted above, we looked into pairwise MAs, but we never had data from at least three similar studies for any of the comparisons.
Peer Reviewer #1	Results	Regardless of pairwise MA, it would be helpful to graphically show the effect sizes in a forest like plot for various comparisons. The text did a good job to explain the results from individual studies and provide estimates of effect size in most sections, but it is hard to get a general and clear idea from the text.	The point of the reviewer is well taken. Forest plots are a very good way to depict effect estimates. As mentioned earlier, we have divided our original plots into smaller subplots about specific comparisons for each outcome to present effect sizes more clearly.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	Page 14 lines 3-4 Did prior treatment use varied by disease duration?	Although this is an interesting question, we did not abstract study data in a way that allows us to answer it.
Peer Reviewer #1	Results	Page 18 line 35 vs. lines 51-52 – same information but the section titles indicated different comparisons.	Thank you, text edited to correctly describe the comparisons evaluated in each section.
Peer Reviewer #1	Results	Page 61, lines 12-14, the observational studies were limited due to medium to high risk of Bias?	We agree the sentence does not make sense, so we have removed it.
Peer Reviewer #1	Results	Page 70: please make sure the description of results to be accurate and consistent. For example, Lines 19-20 "... combination group experienced higher rates of elevated liver enzymes than did patients in the ADA plus MTX group (8.0% vs. 5.0%; p=NR)."	We have corrected this to 8.4% vs 4.0%, p=NR. We also reviewed the results and noted 2 studies (IMPROVED and SWEFOT) were not in our tables at the end of KQ3, and we have updated this.
Peer Reviewer #1	Results	Then Lines 25-26: "... and discontinuation attributable to adverse events (7.8% vs. 10.8%; p=NR) were similar." While the event rates and the difference in event rates were similar in the two cases. "Higher" was used to describe both significant and insignificant differences, too (e.g., line 48).	We have made sure to note when significance was reported or not. We only note that a result is significant if the authors describe it as such. Otherwise, we have adjusted the text to report numbers and let the reader decide.
Peer Reviewer #1	Results	Page 80 lines 13-24 Would be helpful to report the actual effect size.	The adalimumab trial (HOPEFUL 1) does not present any effect size (it is a regression analysis). For the etanercept trial (Enbrel ERA), we already present all of the information relevant to KQ4.
Peer Reviewer #2	Results	In an encyclopedic review such as this one, it is challenging to decide how to, and how much to, present in the Results section. I think the authors have achieved a nice balance here. The text is rather dry reading, while the tables are easier to digest, but both are important.	Thank you and noted.
Peer Reviewer #2	Results	One thing that concerns me is that the authors may not have carefully distinguished studies in which MTX (or other csDMARD) was new therapy vs those in whom it was background therapy. Some studies in early RA focus on those naive to treatment and, in these cases, it is fair to say that MTX monotherapy was compared to the treatment in the other arm(s). However many studies in 'early' RA focus on MTX (or other csDMARD) inadequate responders, who are	Yes, our NWMA was limited to treatment naïve patients. We have gone through the results and grouped the treatment naïve and resistant groups qualitatively.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>then randomized to a new therapy vs placebo (e.g., MTX + Enbrel vs MTX + Placebo). In this case, it is conceptually inaccurate to say that combination was better than monotherapy, but more accurate to say that treatment 2 was better than placebo, in MTX inadequate responders. Instead the authors appear to have considered these two different designs as one group which may have promoted their conclusion that combination therapy (MTX + biologic) is better for initial therapy than monotherapy. Studies confirm that the percent of inadequate responders to MTX who respond to subsequent biologic therapies is less in general than for treatment-naïve patients. I do see, however, as I re-look at the methods, that this distinction was factored into the network meta-analyses which is reassuring.</p>	
Peer Reviewer #2	Results	<p>Another issue is the utilization of data from secondary outcomes instead of primary outcomes. For example, most trials utilize responses at 6 or 12 months as the primary outcomes, while data at 2 year or longer time points are usually secondary outcomes. By the time of these longer followup endpoints, there has been attrition and cross-over from one treatment arm to another. While these longer time points are better for evaluation of safety given the longer exposures, I am surprised they were used for efficacy evaluations. For example, on page 33, at the end of the paragraph labeled 'Etanercept' and in the corresponding table 7 (page 43), the ACR 20 responses from the Enbrel Early RA (ERA) trial indicate statistically higher response for the etanercept group compared to MTX, but the primary outcome was at 12 month and responses at 12 months were not significantly different. Another example is on page 35 at the end of the paragraph labeled 'Rituximab' where again 24 month data are presented, despite that the primary outcome was at 12 months. There are other examples but I did not catalog them but, in general, this approach should be explained and/or defended (unless I missed it).</p>	<p>Thank you for catching these. We agree and have modified our text to note when the outcome timepoint is secondary. For example:</p> <p>U-Act Early: primary outcome was 24 weeks (pg. 41);</p> <p>Enbrel ERA: primary outcome was 12 months (pg. 45);</p> <p>IMAGE trial: we modified our text to focus on the 12-month results (pg. 48);</p> <p>PROWD study: primary outcome was work disability (pg. 68);</p> <p>ASPIRE trial: work disability was secondary outcome measure (pg. 69)</p>
Peer Reviewer #2	Results	<p>By the way, in the above paragraph on page 35 about Rituximab, there is also an error stating that one of the</p>	<p>Thanks for catching this; we have also adjusted the RTX plus MTX arm accordingly.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		treatment arms was rituximab monotherapy. Rituximab is not approved by the FDA for RA as monotherapy, and in this trial, the other arm was not rituximab monotherapy but rather rituximab 500 mg + MTX.	
Peer Reviewer #2	Results	Finally, I found the analyses and summary of findings for the csDMARDs particularly helpful. These are tough studies to compare given the number of them, the variability of treatment regimens, the confounding effect of steroids, etc, so these analyses are a great addition to what's out there.	Thank you.
Peer Reviewer #3	Results	Across all results, it would be ideal to ultimately have enough evidence to draw clear distinctions on effectiveness between therapies, particularly if they can be individualized to subpopulations, but the authors made it clear that this is not possible at this time.	Unfortunately, not possible at this time.
Peer Reviewer #3	Results	This is both a "results" and "methods" comment about the repeat finding that adults over age 65 have less benefit from certain treatments and higher risk of SAEs (multiple references, but page 79, lines 41-43 most clearly states it). My concern from a patient engagement standpoint is in the adverse consequences that may result from findings such as this and the need to ensure that "baselines" are set appropriately to account for smaller intervals of benefit that could occur in this subset of patients. (If they already have some permanent disability they shouldn't be evaluated by the same metrics as a 25 year old newly diagnosed patient). Were there risk adjustment strategies employed or some other mechanism used to allow for elasticity here?	This result is based on a single retrospective subgroup analysis that did not adjust for any baseline risks. Consequently, we have little confidence in these results and graded the strength of evidence as insufficient.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Page 21, KQ1, Non-TNF Biologic Versus Either Non-TNF Biologic or MTX <u>Recommendation:</u> Update text "Three RCTs compared corticosteroids versus csDMARD monotherapy" to read "Non-TNF Biologic versus either Non-TNF Biologic or MTX" <u>Rationale:</u> This section is on non-TNF biologics, therefore text needs correcting.	After further review, we revised this text so that it now reads as recommended but with some modification: "Non-TNF Biologic alone or combined with MTX versus MTX monotherapy".

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
<p>Public Commenter #4 (Leticia Ferri on behalf of BMS)</p>	<p>Results</p>	<p>Page 30, KQ1, Figure 6: Forest plot for NMA: change from baseline in radiographic joint damage score</p> <p><u>Recommendation</u>: Please consider adjusting for heterogeneity across studies at baseline study-level (including radiographic joint damage) via meta-regression.</p> <p><u>Rationale</u>: It is well established that bone erosions in RA patients are associated with baseline disease activity, acute phase reactants (CRP in particular) and baseline erosions. The trials included in this NMA differ in their baseline characteristics, and thus it is important to control for these factors in a meta-regression analysis or other statistical methods. Predisposition of individuals with early RA with existing joint erosion (independent of synovitis) to development of joint space narrowing, which leads to more joint erosion, has been well established via double blind RCTs [1]. This heterogeneity is observed across the trials included for analysis for the outcome of change from baseline in radiographic joint damage score (erosion score, and joint space narrowing). Similarly, heterogeneity can also be seen across included trials for RF positivity and functional disability which are listed as poor prognostic factors for early RA (see uploaded summary table). Also, duration of RA is reported differently across the included trials where some of the trials have reported it as “duration of symptoms” while others have not defined whether the reported value is from initiation of symptoms or diagnosis by treating physician (which could potentially be 2 different time points making the selection of studies for this review with the specific definition of early RA difficult and partly subjective). These systematic differences in baseline patient and study characteristics can have an impact on the treatment effect (effect modifiers) across the different treatment comparisons in the network [2].</p> <p>Additionally, for assessment of radiologic changes, not all included trials reported the measurement tool used (e.g. some trials reported using the Sharp score with different ranges while others did not specify which tool they used). Since the</p>	<p>We used random effects models for the NWMAs, which adjust for between-study heterogeneity. We employed strict eligibility criteria for studies that we included for the NWMA to meet the transitivity assumption: (1) patients with early RA had not failed a prior treatment attempt 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.</p> <p>Therefore, we are confident that clinical heterogeneity across studies should not be problematic for our analyses.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018

Commentator & Affiliation	Section	Comment	Response
		<p>degree of erosion is pre-determinant of comparability of efficacy across trials, the authors should have accounted for this baseline difference across all trials by implementing appropriate statistical methods or described these limitations in detail in the report. Additionally, the newer trials include digital X-rays and these should not be mixed with the earlier study that had standard X-rays.</p>	
<p>Public Commenter #4 (Leticia Ferri on behalf of BMS)</p>	<p>Results</p>	<ul style="list-style-type: none"> • Page 27: Figure 6: Forest plot for network meta-analysis: ACR50 response rate • Page 30: Figure 6: Forest plots for network meta-analysis: change from baseline in radiographic joint damage score • Page 34, Table 14: Summary of the 30 studies included in mixed treatment comparison meta-analysis [3] (Previous SLR report) <p><u>Recommendation:</u> The authors did not separate the analyses by timepoint for each of the outcomes of interest and the outcome measure was assessed at the different timepoints across trials. Given most of the studies reported data at 24 weeks (abatacept trials and others), performing an NMA at 24-week would have been recommended. For other timepoints a sensitivity analysis could have been performed.</p> <p><u>Rationale:</u> Because time may be an important effect modifier, combining all time points may introduce confounding. An NMA of observed relative treatment effects at different timepoint is prone to bias indirect estimates relative to interventions and the outcomes of interest. A different timepoint was used in both analysis regarding ACR and radiographic joint damage score. Thus, the analyses of 24 weeks data only should be considered for the primary analysis as suggested by the NICE report [4], and the analysis including all timepoint should be considered secondary.</p>	<p>We separated the analyses by timepoint. 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 timepoints, data were insufficient for NWMA, or clinical heterogeneity across trials was too high to derive meaningful estimates from NWMA.</p>
<p>Public Commenter #4 (Leticia Ferri on</p>	<p>Results</p>	<p>Page 27: Results section</p> <p><u>Recommendation:</u> the results of the evidence synthesis is not clearly presented and there is no distinction as to which trial is</p>	<p>We have added tables to the report that list all studies used in our NWMA for each outcome of interest.</p>



Commentator & Affiliation	Section	Comment	Response
behalf of BMS)		used for the forest plots. Please present the data which have been included in each analysis. <u>Rationale:</u> As described in NICE checklist [5] for presenting the results of NMA, the table and outcomes did not show which data have been included in each of the analysis.	
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Page 34-35, KQ1, Abatacept <u>Recommendation:</u> Make it clear that the AGREE trial compared ABA + MTX to Placebo + MTX. <u>Rationale:</u> Comparator group received placebo + MTX.	Noted and added.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Page 46, KQ2, Key Points <u>Recommendation:</u> The Key Points for KQ2 should include statement on non-TNF biologics <u>Rationale:</u> On page 51 and in Table 8 non-TNF biologics are included and efficacy of these agents on patient reported outcomes of functional status and quality of life summarized and thus this should be included in the key points. By omitting non-TNF inhibitors from key points, it may be perceived that the non-TNFs do not have data on these critical outcomes	Thank you for the feedback. We added a sentence to include non-TNF biologics to the Key Points.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Page 56, KQ2 Table 8 <u>Recommendation:</u> Add sub-heading in table “Non-TNF Biologics” <u>Rationale:</u> Currently the table includes non-TNF biologics, but there is no sub-heading.	Noted and updated.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Page 60, KQ3 <u>Recommendation:</u> For Table 9, in the AGREE trial, the data for “serious adverse events” in the results column should be corrected to “7.8% and 7.9%” [6] <u>Rationale:</u> Inaccurate representation of the data since the 1.2% vs. 1.2% are discontinuation rates due to serious adverse events, not the percentage of serious adverse events.	Thank you, we have corrected this.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	<p>Pages 61 to 79 Detailed Synthesis, Discussion p. 85</p> <p><u>Recommendation</u>: Conduct NWMA for serious adverse events or explain more precisely why it was not performed.</p> <p><u>Rationale</u>: Statements are made about there not being differences in serious adverse events between the groups, but no data comparing the groups were reported in the narrative summaries for each treatment comparison. Statements were also made in the Discussion about there not being any significant differences in serious adverse events, even though no analyses were reported for the comparisons. NWMA was provided for withdrawals and withdrawals due to adverse events, but not for serious adverse events. Analysis (or a full narrative summary) of the safety outcomes is an important part of every SLR and NMA as it helps bring about balance in assessing relative benefits and harms of the treatment options making a better and more justified case for overall recommendations.</p>	<p>We chose to conduct NWMA of discontinuations due to AEs to account for this.</p> <p>SAEs by themselves were reported qualitatively in this report. We have clarified when we have statistical comparisons, and when we do not, we just present the numbers.</p>
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	<p>Pages 61 to 79 Detailed Synthesis, Discussion p. 85</p> <p><u>Recommendation</u>: Report treatment-related deaths in narrative summaries for each comparison and in the evidence tables.</p> <p><u>Rationale</u>: Deaths are included in the data for serious adverse events in the narrative summary and evidence tables and not reported separately, so it's difficult to tell whether any of the studies reported deaths for specific drugs.</p>	<p>Death is considered a serious adverse event. When reported, deaths were rare, and there were no differences between arms. Details on deaths are noted in the evidence tables.</p>
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	<p>Page 77, KQ3, Table 9: Discontinuation rates and adverse events</p> <p><u>Recommendation</u>: For Table 9, in the AGREE trial, values for upper respiratory infection should be corrected to 10.2% vs 10.3%⁶.</p> <p><u>Rationale</u>: The values currently reported in the draft (26% vs 26%) are the number of patients in each group.</p>	<p>Thank you. We have reviewed and updated this.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	<p>Page 78, KQ3, Table 9: Discontinuation rates and adverse events</p> <p><u>Recommendation</u>: For Table 9, in the AVERT trial, dosing for MTX should be corrected to 7.5-20 mg/week. [7]</p>	Thank you. We have updated this.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	<p>References:</p> <p>[1] Landewe R, Smolen JS, Florentinus S, Chen S, Guerette B, van der Heijde D. Existing joint erosions increase the risk of joint space narrowing independently of clinical synovitis in patients with early rheumatoid arthritis. <i>Arthritis research & therapy</i>. 2015;17:133.</p> <p>[2] Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. <i>Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research</i>. 2014;17(2):157-173.</p> <p>[3] Donahue KE, Jonas DE, Hansen RA, et al. <i>Drug Therapy for Rheumatoid Arthritis in Adults: An Update</i>. Rockville (MD)2012.</p> <p>[4] National Institute for Health and Care Excellence. Sarilumab for moderate to severe rheumatoid arthritis. 2017; https://www.nice.org.uk/guidance/ta485/documents/final-appraisal-determination-document. Accessed 2/2/2018.</p> <p>[5] National Institute for Health and Care Excellence. NICE DSU Technical Support Document 7: Evidence Synthesis of Treatment Efficacy in Decision Making: A Reviewer's Checklist. [Report]. 2012; http://scharr.dept.shef.ac.uk/nicedsu/wpcontent/uploads/sites/7/2016/03/TSD7-reviewer-checklist.final_.08.05.12.pdf. Accessed 2/2/2018.</p>	<p>Noted. See below for details on whether articles used in report:</p> <ol style="list-style-type: none"> 1) YES - PREMIER study companion 2) NO - Not in database, but not used in report because commenter cited it to support their statement. 3) YES - The previous 2012 RA SR 4) NO - Not in database, but not used in report because commenter cited it to support their statement. 5) NO - Not in DB, but not used in report because commenter cited it to support their statement. 6) YES – AGREE study companion. 7) YES – AVERT study parent article.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>[6] Smolen JS, Wollenhaupt J, Gomez-Reino JJ, et al. Attainment and characteristics of clinical remission according to the new ACR-EULAR criteria in abatacept-treated patients with early rheumatoid arthritis: new analyses from the Abatacept study to Gauge Remission and joint damage progression in methotrexate (MTX)-naive patients with Early Erosive rheumatoid arthritis (AGREE). Arthritis research & therapy. 2015;17:157.</p> <p>[7] Emery P, Burmester GR, Bykerk VP, et al. Evaluating drug-free remission with abatacept in early rheumatoid arthritis: results from the phase 3b, multicentre, randomised, active-controlled AVERT study of 24 months, with a 12-month, double-blind treatment period. Ann Rheum Dis. 2015;74(1):19-26.</p>	
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Results	Tables: See Results section.	Responded in the Results section .
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 16</p> <p><u>Comment:</u> The Limitation Section highlights that the evidence on benefits and harms from head-to-head clinical trials in early RA is limited. However, to facilitate transparency and increase readers' comprehension of the limited evidence and subsequent assessment, the Draft Report should:</p> <p>Explicitly state the disparity in the number of studies available per intervention identified from peer-reviewed published literature in the conclusions for each Key Question (and Key Points).</p> <p>Consider carefully presenting in Table 6 data on the specific number of publications identified for each drug, based on the selection criteria, to address the Key Questions and inform final conclusions.</p>	Thank you. We have added the number of studies evaluating each eligible drug to Table 6.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 33</p> <p><u>Comment:</u> The statement on ETN and ACR 50 response from COMET is misleading by including the term “only” in the following statement – “patients in the ETN plus MTX group had a higher ACR 50 response than MTX only at 52 weeks (70.7% vs. 49%; p<0.001).” – The COMET trial reported on the primary endpoint at 52 weeks. No other timepoints were reported. The insertion of “only” suggests multiple timepoints were reported and only the 52 week timepoint was significant. This is inconsistent with the findings of the COMET trial.</p> <p>It should be noted that the coprimary endpoints were assessed at 52 weeks, and findings on ACR 50 response prior to 52 weeks were not assessed.</p>	We have clarified the text. It now reads: patients in the ETN plus MTX group had a higher ACR50 response than MTX monotherapy at 52 weeks.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 33 [continued]</p> <p><u>Comment:</u> DAS remission was incorrectly defined as DAS 28 <1.6 instead of <2.6.</p>	Noted and updated.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 34</p> <p><u>Comment:</u> An incorrect reference has been made to ETN on page 34. This should be revised to read the correct drug.</p>	We reviewed page 34, and the mention of ETN was correct with the NWMA. We have added a subplot below the text to further clarify.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 46</p> <p><u>Comment:</u> As alluded to above in comments on the Evidence Summary regarding the conclusions made in the Draft Report on the effect of ETN on functional capacity, a similar update should be made to the relevant Key Point on page 46 which suggests that only combinations of adalimumab plus MTX and infliximab plus MTX but no other TNF biologics produce statistically significant greater improvements in functional capacity in comparison to MTX alone. Data from the COMET</p>	Noted and clarified in the text of the report. We have added that evidence is inconclusive for the TNF biologic ETN.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		trial reported on page 50 of the Draft Report provides supporting evidence for ETN.	
Public Commenter #5 (Jason Spangler on behalf of Amgen)	Results	<p>Page 79</p> <p><u>Comment:</u> In the Key Points for Key Question 4, the conclusions drawn for ETN and MTX appear to be unduly definitive with use of the phrase “substantially higher risks,” despite the grade assignment of low levels of evidence (LOE), which limit the confidence regarding the true estimate of effect. It is unclear how definitive conclusions can be drawn when the supporting study⁴ used as a basis for the conclusion is considered deficient as conclusions about safety in patients older than 65 years were made based on a very small subpopulation (n=37; 18% of the total study population) and comparative statistical analyses were not completed.</p> <p>It should be explicitly stated that evidence regarding comparator treatments evaluating the potential harms is limited, and as a result, no conclusions can be drawn regarding the relative comparative safety outcomes of ETN versus other TNFs evaluated in this review.</p>	This is a valuable point. After reconsideration, we have downgraded this comparison to insufficient.
Public Commenter #6 (Fang Sun on behalf of Merck)	Results	<u>P23:</u> The authors conclude: “With respect to combination therapy, long-term studies show no differences in remission rates between initial combination versus step-up therapies (moderate SOE).” Please specify what combination therapy is referred here.	Initial combination therapies were several including csDMARDs and TNF biologics, MTX plus SSZ plus PRED group, and the MTX plus IFX group, as noted in the BeSt trial, and MTX plus ETN in the TEAR study. The Key Points refer to the detailed results.
Public Commenter #6 (Fang Sun on behalf of Merck)	Results	<u>P39 Table 7:</u> The proportion of early RA patients enrolled in the studies should be presented in the table.	No studies in Table 7, except one, reported the proportion of early RA patients they enrolled. Please refer to Appendix C (Evidence Tables) for studies’ mean and median disease duration, when available.
Public Commenter #6 (Fang Sun on behalf of Merck)	Results	<u>Table P53:</u> The proportion of early RA patients enrolled in the studies should be presented in the table	No studies in Table 8, except one, reported the proportion of early RA patients they enrolled. Please refer to Appendix C (Evidence Tables) for studies’ mean and median disease duration, when available.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Discussion	Limitations: Some weakness of the NWMA were well discussed. Still the data were too sparse to warrant a formal MA, and the majority of comparisons are only indirect comparisons.	Because data were insufficient for pairwise meta-analyses, we had to rely on NWMA. Results from NWMA have limitations, as we outline in the Discussion.
Peer Reviewer #1	Discussion	Page 81 lines 30-31 “All outcome differences reported in the table are statistically significant.” ?? In Table 10? No outcome differences were reported.	After consideration, we have removed this statement because this table is a summary of multiple studies.
Peer Reviewer #2	Discussion	The Discussion section is overall well written and clear, and doesn't overreach the findings. I do think the caveat that secondary outcome measures were used in some analyses needs to be mentioned, as this may also explain differences in findings between this study and those of the ACR or EULAR.	We have added this caveat.
Peer Reviewer #2	Discussion	Clearly, more comparative effectiveness clinical trials with the biologics need to be done, but whether these need to be directed specifically to early RA patients is a question here since biologics tend not to be first line treatments.	We have added this comment to the ‘Research Needs’ section in the Discussion.
Peer Reviewer #3	Discussion	I think the entire future research needs section (pages 89-90) was well done and clearly outlines the needs of the patient community. I would only add that not only is additional research needed, additional patient-centered research is needed with appropriate use of PROs and other PGHD so that results are truly reflective of patient preferences and desires.	We agree and have added this statement to the ‘Research Needs’ section in the Discussion.
Peer Reviewer #4	Discussion	The discussion needs to be more balanced than it currently is, and needs to be based on the data presented.	We have toned down the Discussion in line with the Evidence Summary. We have noted limitations of the NWMA. We have also added text noting the difficulty of balancing the higher efficacy of combination therapy with a burden of multiple drugs and potentially higher risk of harms, with this as the first-line therapy.
Peer Reviewer #4	Discussion	Clarity and Usability: Most points are clearly presented in results. Please see my comments related to the discussion section. This is an updated analysis and most findings that support previous reviews. With more data and an updated analysis, it advances knowledge and provides a higher quality	We agree and describe our limitations in the Limitations and Future Research Needs section of the Discussion (lack of head-to-head trials, reliance on NWMAs, subgroup needs).

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		evidence for some comparisons. However, most of the important questions still have limited high-quality data to help make clear choices/decisions.	
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Discussion	Page 89, Research Needs <u>Recommendation:</u> Please consider including real-world data, especially registry data, as these studies include data on long term effects and follow-up. <u>Comment:</u> There was no real discussion about the clinical significance of the results (magnitude of effect).	We have added text regarding real world data/registry data in the discussion.
Public Commenter #6 (Fang Sun on behalf of Merck)	Discussion	<u>P88:</u> Patient discontinuation rates varied significantly across the RCTs and observational studies discussed in the first paragraph of the page. Reasons for this variation and how it may impact on the applicability of the results should be discussed.	Yes, as we noted in the Applicability section, discontinuation rates were higher in observational and non-controlled studies. We have added the higher discontinuation rates in observational studies may reflect real world settings as compared with the tighter adherence in a controlled clinical trial
Public Commenter #6 (Fang Sun on behalf of Merck)	Discussion	<u>P88:</u> Discussions on the selected studies' differences and how they may impact on the meta-analysis results should be considered.	We had very strict criteria for studies to be eligible for NWMA. To ensure the transitivity assumption, studies had to meet the following criteria: (1) patients with early RA had not failed a prior treatment attempt 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. Therefore, we believe that the impact of studies' differences on effect estimates is small.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	References	Pages 49 and 50 <u>Comment:</u> A few incorrect references were identified. The ERA study cited as the second eligible study with information on the impact of ETN on physical function (reference 79 in Draft Report) on page 50 does not include results on functional capacity. Although this is the reference for the clinical trial, the correct publication (reference 76 in Draft Report) which presents results reported should be included.	Noted and updated.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		The above reference should be corrected on page 49.	
Public Commenter #5 (Jason Spangler on behalf of Amgen)	References	<u>Comment:</u> Reference 86 in the Draft Report has been wrongly cited as an ETN study. This publication is on adalimumab.	Thank you. We found one incorrect use of reference 86 with the Enbrel ERA, and we have removed that from the text.
Public Commenter #7 (Tom Innal on behalf of Genentech)	References	<ul style="list-style-type: none"> Consider incorporating data from the following references for Actemra to ensure the Network Meta-Analysis is comprehensive. <p><u>FUNCTION</u></p> <p>Burmester GR, Rigby WF, van Vollenhoven RF, et al. Tocilizumab combination therapy or monotherapy or methotrexate monotherapy in methotrexate-naïve patients with early rheumatoid arthritis: 2-year clinical and radiographic results from the randomised, placebo-controlled FUNCTION trial. <i>Ann Rheum Dis</i>. 2017 Jul;76(7):1279-1284. https://www.ncbi.nlm.nih.gov/pubmed/28389552</p> <p>Burmester G, Blanco R, Keiserman M, et al. Tocilizumab as combination therapy and as monotherapy vs methotrexate in MTX-naïve patients with early rheumatoid arthritis: Patient-reported outcomes from a randomized, placebo-controlled trial. <i>Ann Rheum Dis</i> 2014;73(suppl 2):672. Presented at: European League Against Rheumatism 2014; Paris, France. Abstract #SAT0226. http://www.eular.org</p> <p><u>U-ACT-EARLY</u></p> <p>Teitsma XM, Jacobs JW, Welsing PM, et al. Sustained Drug Free Remission in Early RA Patients Treated To Target with Tocilizumab, Methotrexate or Their Combination. Presented at: European League Against Rheumatism 2016; London, UK. Abstract #FRI0203. http://www.eular.org</p> <p>Teitsma XM, Jacobs JW, Welsing PMJ, et al. Effects of tocilizumab in DMARD-naïve early rheumatoid arthritis</p>	<p><u>FUNCTION publications</u></p> <p>Burmester GR, Rigby WF, van Vollenhoven RF, et al.: Retrieved with 10/5/17 update searches and already incorporated into revised report.</p> <p>Burmester G, Blanco R, Keiserman M, et al.: Found through handsearching and excluded as abstract-only record.</p> <p><u>U-Act-Early publications</u></p> <p>Teitsma XM, Jacobs JW, Welsing PM, et al Sustained Drug Free Remission in Early RA Patients: New record not suggested or found previously, but we cannot include it because it is an abstract-only record.</p> <p>Teitsma XM, Jacobs JW, Welsing PMJ, et al. Effects of tocilizumab in DMARD-naïve early rheumatoid arthritis patients: New record not suggested or found previously, but we cannot include it because it is an abstract-only record.</p> <p>Teitsma XM, Jacobs JW, Welsing PM, et al. Tocilizumab Inhibits Progression of Erosive Joint Damage in Early Rheumatoid Arthritis: New record not suggested or found previously, but we cannot include it because it is an abstract-only record.</p> <p>Teitsma XM, Jacobs JW, Welsing PMJ, et al. Patient-reported outcomes in newly diagnosed early rheumatoid</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>patients on Health-Related quality of life: results of the U-Act-Early trial. Presented at: European League Against Rheumatism 2017; Madrid, Spain; Abstract #SAT0218 http://www.eular.org</p> <p>Teitsma XM, Jacobs JW, Welsing PM, et al. Tocilizumab Inhibits Progression of Erosive Joint Damage in Early Rheumatoid Arthritis More Effectively Than Step-up Methotrexate Therapy. Presented at the 2017 American College of Rheumatology Annual Meeting in San Diego, CA; November 3–8, 2017. ACR Abstract #2458. https://www.rheumatology.org/Annual-Meeting</p> <p>Teitsma XM, Jacobs JW, Welsing PM, et al. Patient-reported outcomes in newly diagnosed early rheumatoid arthritis patients treated to target with a tocilizumab- or methotrexate-based strategy. Rheumatology (Oxford). E-pub Date: [published online ahead of print] September 2017. https://www.ncbi.nlm.nih.gov/pubmed/29029185</p>	<p>arthritis patients: New record with eligible data that we have incorporated into the report.</p>
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Appendix	<p>Appendix A: Search Strings</p> <p><u>Comment 1:</u> There is a significant different in the numbers reported in search tables for the total number of hits for all databases search and PRISMA flow presented as Figure 2 on page 14 of the report. This is lack of clarity and an error in reporting which is concerning especially as this report will become public and it is not known which search (database on citations) were used for the purpose of screening.</p>	<p>We have fixed this by improving transparency in two ways:</p> <ol style="list-style-type: none"> 1) Reporting N's of duplicates removed from each database's initial yield in Appendix A; 2) Referring readers in the Search Results text of the Results to Appendix A for details about duplicates removed
Public Commenter #4 (Leticia Ferri on behalf of BMS)	Appendix	<p>Appendix A: Search Strings</p> <p><u>Comment 2:</u> The list of intervention names used in the searches focuses only on generic names of the drugs while to make the search highly sensitive it would have made more sense to use generic names as well as brand names. The strings will search mainly in the title and abstracts and there might be cases where the authors of the papers have avoided using generic names (and instead had used brand names) throughout the abstract. Additionally, drug classes were not used as part of the search string. This could potentially lead to</p>	<p>Our team considered this, but in peer-reviewed literature, it is generally not allowable to use the drug brand name in the paper. We had also decided to model our search strategy based on our prior report searches.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		missing even more trials as there might have been cases where authors had discussed and mentioned only class comparison in the title and abstract (to be captured only if class names were among the search strings) but presented more detailed information in results tables of the fulltext on individual therapies.	
Peer Reviewer #2	General	This is an ambitious and comprehensive analysis of the comparative effectiveness of RA treatments, and the authors are to be congratulated for a job well done.	Thank you.
Peer Reviewer #2	General	<p>To some extent, arguing over a temporal definition of what constitutes early vs later disease is increasingly anachronistic. Since we know that autoimmune diseases evolve as a continuum starting years before the first clinical symptoms appear, the ascertainment of time of onset of disease has itself taken on new meaning, leading to a relaxation of the classification criteria for RA (which, though not meant to be diagnostic criteria, are nonetheless used as such.)</p> <p>In step with this, our definition of 'early' RA has become more and more restricted, some folks arguing that early RA includes only the first six weeks of symptoms. A number of us have proposed to define established RA as already treated disease, or - specifically for RCTs - that which had failed or inadequately responded to the initial treatment (usually with a csDMARD as monotherapy) while early disease is treatment-naïve disease. In this ARHQ review, I find the choice to restrict this treatment review to early disease somewhat dissatisfying and scientifically unjustified. That is probably my only issue with the overall approach.</p> <p>I would challenge you in your introduction to at least acknowledge the artificiality of defining an early vs late phase in the continuum that starts pre clinically, and justify more clearly why you chose to separate the two (early vs late). I think we would all agree that it is easier to get control of RA in its earlier stage than later but I don't think there are data suggesting that responses to specific therapeutic agents differ by stage of disease, but the unsophisticated reader</p>	We agree with the issues of defining early RA and have added context to the reasoning behind our definition.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		might assume there are in the absence of a scientific justification for splitting out early disease. Therefore, a scientific rationale for choosing to separate by time - other than because the data are overwhelming - would be important to lay out up front.	
Peer Reviewer #2	General	In addition to a philosophical objection to the restriction by time is a practical problem. By one year into disease, the less responsive RA patient has typically already gone through several iterations of treatment and cannot be considered the same as the newly diagnosed treatment-naïve RA patient. To combine them as if they were the same is somewhat unsatisfying. Just as the evolution of RA is a continuum, so is the treatment. I realize it's too late to reconsider early vs late, and the definition of early disease, but acknowledging some of these nuances would be helpful for readers.	We have acknowledged these limitations of the definition in the Introduction and the 'Limitations' section of the Discussion.
Peer Reviewer #2	General	Clarity and Usability: Yes, to all of the above questions. I guess the big issue for us in the field with this report is if, and how, the highlighting that biologic + MTX is better than MTX alone will be used by the manufacturers to pester providers, insurance companies, and the FDA to loosen restrictions about early use of biologics. The feeling out there is that the differences in responses of treatment-naïve patients to biologic + MTX vs MTX are not large enough to shift initial treatment away from MTX monotherapy. [And, compounding the issue, CMS is planning in its revised reimbursement schemes to penalize doctors for prescribing these expensive medications, as if the physicians have control over pricing.]	Noted and agree with the concerns regarding manufacturers encouraging the use of biologics and pestering providers, insurance companies, and the FDA. In the Discussion, we note a few caveats: '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 biologic (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, there are difficulties with balancing the higher efficacy with a burden of multiple drugs and potentially higher risks.'
Peer Reviewer #3	General	Before I begin, I want to provide brief context for my remarks. Because I am a patient reviewer (both a person with a rheumatologic diagnosis and a professional who works with and for patients with arthritis), my review and comments are focused on the reports' inclusion of patient-centered concepts/principles and potential impacts on patients.	Noted.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	General	The report includes a great deal of information that may be useful to clinicians and patients, but as with any guidance it should be used in conjunction with thorough conversations with patients about their preferences and goals for treatment. In terms of the target population and audience, the challenges of the definition of "early" RA were stated, but another issue that was not touched on (unless I missed it) is the significant challenge of delayed diagnosis. Often an onset is insidious and difficult to quickly diagnosis; even when the symptomology is clear patients may struggle to reach a provider who is able to gather enough relevant clinical data to produce a working diagnosis. Because of this, I wonder how many patients included in the reviewed studies had pre-existing disease for many months or years, regardless of the specific date of diagnosis (and hence, inclusion in the population of "early" RA).	We agree and have added this to the 'Limitations' section of the Discussion.
Peer Reviewer #3	General	I was interested and encouraged by the inclusion of the second "contextual" question; from a patient engagement perspective, taking account of the real situational barriers to care is crucial. Evaluating treatment options for relative effectiveness in the absence of that context is far less useful.	Noted.
Peer Reviewer #3	General	I was confused by the inclusion of patient adherence within key question 3 and will address that more in the methods section.	See response above in the Methods section. We agree that lack of adherence can have other reasons than harms (e.g., lack of efficacy). We viewed lack of adherence as an undesirable event. This is the reason why we present it in the harms chapter.
Peer Reviewer #3	General	Key question 4 was also very well-constructed and important from a patient engagement perspective--how patients make decisions about their care and what outcomes are important to them are so heavily influenced by these types of societal and demographic factors, that this context is absolutely critical.	Thank you.
Peer Reviewer #3	General	While the focus of this review is necessarily on medications, we know from the patient community that treatments like massage, physical and occupational therapy, meditation, and other complementary therapies can be incredibly useful and	Noted.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		beneficial and it would be wonderful to have the endorsement of the scientific community on those activities and a better understanding of their effectiveness when included in a holistic treatment approach.	
Peer Reviewer #3	General	Clarity and Usability: Depending on the intended audience, the report reads very well and summarizes an immense amount of information in a clear and concise way. If the intention is for a patient audience to ultimately be able to access the information, I suggest working with a patient advocacy organization, patient advisory team, or both, to produce a revised, shorter, more linguistically appropriate version.	Noted.
Peer Reviewer #4	General	The report is meaningful. The target population is defined reasonably well. The key questions are appropriate and explicitly stated.	Thank you.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	As also described earlier in the comments made for the introduction section, eliminating the search for main conferences and proceedings (such as ACR and EULAR) published during the past 2-3 years in this review might have led to also missing trial reports on stratified population that are a potential match for the specific definition criteria for early RA in this review. One such example of missing evidence in this review is a poster presented at EULAR conference 2014 [1] which stratifies patients with less than or greater than 6 months of disease duration and discusses clinical outcomes in patient with early RA using data from the AMPLE trial [2] (head-to-head trial).	We searched the gray literature, which also included abstracts. Abstracts were excluded given high risk of bias and/or limited data.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	If definition of early RA is expanded to ≤ 2 years, please consider including AMPLE trial.	We considered the AMPLE trial for inclusion previously when BMS sent us supplemental evidence and data last July (see Schiff et al., 2016 ⁸), but we excluded the subgroup analysis data that BMS sent us because of wrong design (post-hoc analysis of subgroups not assigned to treatments by randomization). The original study publication, Schiff et al., 2014 ⁹ , was ineligible because patients with RA up to 2 years were eligible.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	Please consider highlighting that each non-TNF biologic works on a specific target of the autoimmune inflammatory cascade. Among the non-TNF biologics, abatacept is the only bDMARD that has selective T-cell costimulation modulation as its mechanism of action.	We considered this but determined that this was outside the scope of the report, since we do not discuss the mechanism of action for other drugs.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	The corrections to the adverse events noted above suggest that there are issues with the quality of the data extraction – please consider checking again for accuracy of both extraction and reporting.	Thank you, we have rechecked the extraction and data tables.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	With regard to Grading the Strength of Evidence for Major Comparisons and Outcomes; we note that SOE was graded using the EPC approach, but it should be noted that the GRADE working group have updated their guidance for NMA, and this could be considered for a revision to this review. Please see the recent update by Brignardello-Petersen R, Bonner A, Alexander PE, et al [3].	Because the current EPC guidance does not offer much detail about grading strength of evidence for NWMA, we used the new GRADE guidance for data based on NWMA. In general, the principal approaches of GRADE and EPC guidance towards strength of evidence are very similar.
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	<p>SR Report states, “Fourteen studies (14/46= 30%) specifically enrolled only RA patients without a prior treatment history, and six studies [2-7] (13%) did not report any information at all about prior treatment use. Among the remaining 26 studies, 9 studies (35%) reported on patients’ previous use of methotrexate (MTX) use specifically, 15 studies (58%) on any DMARD use, and 14 studies (54%) on corticosteroid use.” [4] At least one other NMA used separate networks for MTX naïve and MTX-IR (inadequate response) studies – given small number of studies, it may not be possible to do this, but there should be some discussion about this.</p> <p><u>Recommendation:</u> Add table under each network with list of studies included in network for each treatment comparison.</p> <p><u>Rationale:</u> Not clear which studies are included in each network analysis. It’s also hard to determine whether baseline</p>	As noted in response to an earlier comment, we have added tables to the report that list all studies used in our NWMA for each outcome of interest.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>
 Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		characteristics impacted the results since it's not clear which studies were included in each network.	
Public Commenter #4 (Leticia Ferri on behalf of BMS)	General	<p>References:</p> <p>[1] Schiff M, Weinblatt ME, Valente R, et al. Clinical response by baseline RA disease duration in the AMPLE (abatacept versus adalimumab comparison in biologic-naïve RA patients with background methotrexate) trial. Poster presented at: EULAR Annual European Congress of Rheumatology; June 11-14, 2017; Paris, France. FRI0019.</p> <p>[2] Schiff M, Weinblatt ME, Valente R, et al. Reductions in disease activity in the AMPLE trial: clinical response by baseline disease duration. RMD open. 2016;2(1):e000210.</p> <p>[3] Brignardello-Petersen R, Bonner A, Alexander PE, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. Journal of clinical epidemiology. 2018;93:36-44.</p> <p>[4] Agency for Healthcare Research and Quality P-CORI. Drug Therapy for Early Rheumatoid Arthritis: A Systematic Review Update. 2017; https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/draft-report-drug-therapy-for-earlyrheumatoid-arthritis.pdf. Accessed 2/2/2018.</p>	<p>We scanned the references to determine if we had already used them, and if so, how. Below are the details:</p> <ol style="list-style-type: none"> 1) Abstract-only 2) Excluded for ineligible study design 3) We are adding this to the Methods 4) Our ongoing RA report
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	<p>GENERAL COMMENTS</p> <p>Thank you for the opportunity to comment on the draft report on “Drug Therapy for Early Rheumatoid Arthritis in Adults: A Systematic Review Update” (herein referred to as “Draft Report”).</p> <p>Amgen is invested in the scientific development of therapies to meet unmet medical needs and enrich patients’ lives, and appreciates the efforts of the Agency for Healthcare Research and Quality (herein referred to as “Agency”) in developing this evidence report with the aim of helping healthcare stakeholders make informed decisions and improve quality of care for patients with early rheumatoid arthritis (RA).</p>	Noted.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		Amgen manufactures etanercept (Enbrel®) (herein referred to as “ETN”), a tumor necrosis factor (TNF) receptor, which is approved by the Food and Drug Administration (FDA) “for reducing the signs and symptoms, inducing major clinical response, inhibiting the progression of structural damage, and improving physical function in patients with moderately to severely active rheumatoid arthritis.” ¹	
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	Amgen believes that evidence from systematic reviews should be reported accurately, comprehensively, and in a transparent manner to ensure that stakeholder assessments are based on a solid evidence base. Amgen acknowledges that the authors of the Draft Report have strived for very high quality in conducting the systematic review for a complex early RA patient as alluded to in the methods utilized for the systematic review and guidelines referenced (Agency’s Methods Guide for Effectiveness and Comparative Effectiveness Reviews and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)). To further improve the accuracy of reporting and reliability of the evidence, Amgen is providing comments on outcome evidence relating to these patients and the drugs, including ETN, summarized in this report. These outcomes include disease activity, functional capacity, and harms with specific examples mentioned in the attached comments.	Noted.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	<p>The summary of evidence across treatment interventions is lacking appropriate interpretation and balance. For example, when summarizing evidence on benefits for biologic disease modifying anti-rheumatic drug (bDMARDs), only two drugs have been called out, where other TNFs, including ETN, with equivalent supporting evidence on benefits are not.</p> <p>To facilitate assessments of strengths and weaknesses of evidence, conclusions should accurately represent findings in the context of the totality of evidence. Certain conclusions appear to be based on limited scientific evidence. The Draft Report did not consider the strength of available evidence and the paucity of evidence on specific outcomes in areas such as</p>	After further review, we have broken apart the large forest plots and present each subplot of comparison with each section. We then modified the text in each of the comparison paragraphs to describe the relevant comparisons (significant and nonsignificant).

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>
 Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		harms of these drugs. Specific instances are outlined in the attached comments. Additionally, a few errors in reporting were identified.	
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	In defining the scope of the Draft Report, the Agency has modified the scope of the Draft Report to include only those patients with early RA. Given this targeted scope of the Draft Report, there are substantial differences in the body of evidence and conclusions in comparison with the 2012 Report, as such the Draft Report is only a partial update. The Agency should be clear this is a partial update of the 2012 Report.	We modified the text to make this clearer in the beginning of the report.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	Amgen looks forward to the Agency's response and changes to this important document to ensure the final version is accurate and appropriately addresses the evidence for the treatment of early RA patients.	Noted.
Public Commenter #5 (Jason Spangler on behalf of Amgen)	General	<p>Dear Reviewer, Amgen is providing you with referenced information. If you would like a reprint of a reference, contact Amgen Medical Information. Please note that if you are a covered recipient as defined by the Affordable Care Act (ACA) and Amgen provides you with the requested reprint(s), Amgen's cost to obtain such reprint(s) may need to be disclosed and reported in accordance with the requirements under the ACA, state law and related disclosure obligations by Amgen. If you are a non-covered recipient requesting information on behalf of or the benefit of a covered recipient (physician or teaching hospital), the same requirements may apply.</p> <p>Again, thank you for the opportunity to provide comments on the draft report on "Drug Therapy for Early Rheumatoid Arthritis in Adults: A Systematic Review Update." Please contact Jason Spangler or George Jaresko with any questions regarding this submission or requests for further information.</p>	Noted.
Public Commenter	General	Currently, there is a lack of agreement on the definition of early RA. As suggested by this draft review, most existing	As noted in above, we chose this narrower scope due to stakeholder and expert input.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
#6 (Fang Sun on behalf of Merck)		<p>studies reported outcomes for a mixed population including all stages of the disease. Due to the significant flaw in study selection (see previous comments), the findings of this draft review may not be generalizable to the early RA population.</p> <p>We suggest PCORI adhere to the original scope of the 2012 report for this update, which would yield more helpful information to guide RA management. It would be more meaningful to include all stage RA patients in the review because RA is a continuous disease.</p>	
Public Commenter #6 (Fang Sun on behalf of Merck)	General	<p>Biosimilars are a cost-effective treatment option that recently became available to RA patients. Patients, healthcare providers, and payers need meaningful and timely information on the effectiveness and safety of biosimilars. In the case where no study was identified that could meet the CER's stringent inclusion criteria, the authors should include other information or studies for discussion or analysis. Filling the knowledge gap on the effectiveness and safety of biosimilars would make this review more valuable, useful to the public and other stakeholders.</p>	<p>We appreciate this feedback. Our report had stringent inclusion criteria for studies and our scope was for early RA patients only. This should be a focus of future studies.</p>
Public Commenter #7 (Tom Innal on behalf of Genentech)	General	<p>Dear AHRQ Clinical Expert Panel:</p> <p>Genentech appreciates the opportunity to provide comments on the Agency for Healthcare and Quality (AHRQ) Systematic Review for Drug Therapy in Early Rheumatoid Arthritis (RA). We applaud AHRQ's efforts in facilitating an open source, transparent process for providing this clinical source of healthcare technologies from a broad set of stakeholders.</p> <p>Genentech is committed to advancing the scientific understanding of RA and pursuing the development of novel therapies to help individuals diagnosed with this chronic and debilitating disease. Early and appropriate treatment by rheumatologists can help improve outcomes and prevent unnecessary delay in determining appropriate drug regimens for patients. Treatment decisions for patients with RA are complex and personal. We support comparative effectiveness tools that account for the needs of individual patients, facilitate</p>	<p>Noted.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018



Commentator & Affiliation	Section	Comment	Response
		<p>a meaningful dialogue between patients and providers, and advance conversations around clinical value.</p> <p>Our comments are focused on patient characteristics for the systematic review, results from Actemra's clinical trials in early RA and additional data for Actemra which should be considered in the systematic review. Although our comments primarily address the Executive Summary Section, these should be incorporated throughout the report. Please see the below recommendations that will help improve the accuracy and applicability of the evidence report AHRQ has generated.</p>	
Peer Reviewer #1	Quality of the Report	N/A	N/A
Peer Reviewer #2	Quality of the Report	Superior	Thank you.
Peer Reviewer #3	Quality of the Report	Good	Thank you.
Peer Reviewer #4	Quality of the Report	Good	Thank you.

Source: <https://effectivehealthcare.ahrq.gov/topics/rheumatoid-arthritis-medicine-update/final-report-update-2018>

Published Online: July 16, 2018

References

1. Peterfy CG; Wu C; Szechinski J; Dicarolo JC; Lu Y; Genovese M; Strand V; Genant HK. Comparison of the Genant-modified Sharp and van der Heijde-modified Sharp scoring methods for radiographic assessment in rheumatoid arthritis. *Int J Clin Rheumatol*. 2011;6(1):15-24. doi: 10.2217/ijr.11.1.
2. Burmester GR; Rigby WF; Van Vollenhoven RF; Kay J; Rubbert-Roth A; Kelman A; Dimonaco S; Mitchell N. Tocilizumab in early progressive rheumatoid arthritis: FUNCTION, a randomised controlled trial. *Ann Rheum Dis*. 2016;75(6):1081-91. doi: 10.1136/annrheumdis-2015-207628.
3. Bijlsma JW; Welsing PM; Woodworth TG; Middelink LM; Petho-Schramm A; Bernasconi C; Borm ME; Wortel CH; ter Borg EJ; Jahangier ZN; van der Laan WH; Bruyn GA; Baudoin P; Wijngaarden S; Vos PA; Bos R; Starmans MJ; Griep EN; Griep-Wentink JR; Allaart CF; Heurkens AH; Teitsma XM; Tekstra J; Marijnissen AC; Lafeber FP; Jacobs JW. Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. *Lancet*. 2016 Jul 23;388(10042):343-55. doi: 10.1016/s0140-6736(16)30363-4. PMID: 27287832.
4. Ren L; Guo P; Sun QM; Liu H; Chen Y; Huang Y; Cai XJ. Number of parity and the risk of rheumatoid arthritis in women: A dose-response meta-analysis of observational studies. *J Obstet Gynaecol Res*. 2017 Sep;43(9):1428-40. doi: 10.1111/jog.13370. PMID: 28613016.
5. Singh JA; Saag KG; Bridges SL, Jr.; Akl EA; Bannuru RR; Sullivan MC; Vaysbrot E; McNaughton C; Osani M; Shmerling RH; Curtis JR; Furst DE; Parks D; Kavanaugh A; O'Dell J; King C; Leong A; Matteson EL; Schousboe JT; Drevlow B; Ginsberg S; Grober J; St Clair EW; Tindall E; Miller AS; McAlindon T; American College of Rheumatology. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2016 Jan;68(1):1-25. doi: 10.1002/acr.22783. PMID: 26545825.
6. Ioannidis JP; Evans SJ; Gotzsche PC; O'Neill RT; Altman DG; Schulz K; Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004 Nov 16;141(10):781-8. PMID: 15545678.
7. Durez P; Malghem J; Nzeuseu Toukap A; Depresseux G; Lauwerys BR; Westhovens R; Luyten FP; Corluy L; Houssiau FA; Verschueren P. Treatment of early rheumatoid arthritis: a randomized magnetic resonance imaging study comparing the effects of methotrexate alone, methotrexate in combination with infliximab, and methotrexate in combination with intravenous pulse methylprednisolone. *Arthritis Rheum*; 2007. p. 3919-27.
8. Schiff M; Weinblatt ME; Valente R; Citera G; Maldonado M; Massarotti E; Yazici Y; Fleischmann R. Reductions in disease activity in the AMPLE trial: clinical response by baseline disease duration. *RMD Open*. 2016;2(1):e000210. doi: 10.1136/rmdopen-2015-000210. PMID: 27110385.
9. Schiff M; Weinblatt ME; Valente R; van der Heijde D; Citera G; Elegbe A; Maldonado M; Fleischmann R. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. *Ann Rheum Dis*. 2014 Jan;73(1):86-94. doi: 10.1136/annrheumdis-2013-203843. PMID: 23962455.