Cycling of tumor necrosis factor inhibitors *versus* switching to different mechanism of action therapy in rheumatoid arthritis patients with inadequate response to tumor necrosis factor inhibitors: a Bayesian network meta-analysis

Alberto Migliore iD **, Giuseppe Pompilio** iD **, Davide Integlia, Joe Zhuo and Evo Alemao***

Abstract

Introduction: For patients with rheumatoid arthritis (RA) with an inadequate response to tumor necrosis factor inhibitors (TNFi), main options include cycling onto a different TNFi or switching to a biologic/targeted synthetic disease-modifying antirheumatic drug with a different mechanism of action (MOA). This network meta-analysis (NMA) assessed comparative clinical efficacy of cycling *versus* switching.

Methods: We conducted a literature search in MEDLINE, Embase, and Cochrane Library. Outcomes included proportion of patients with 20%, 50%, or 70% response to American College of Rheumatology criteria (ACR20/ACR50/ACR70 response), Disease Activity Score in 28 joints (DAS28) score below 2.6 or between 2.6 and 3.2, mean change in DAS28 score, mean reduction in and proportion of patients achieving a clinically meaningful reduction ($\geqslant$0.22) in Health Assessment Questionnaire score, number of serious adverse events (AEs), and withdrawals for any reason/due to AEs/lack of treatment efficacy. To account for the wide range of study populations and designs, we developed three models to conduct the NMA: fixed-effect, random-effects, and hierarchical Bayesian. PROSPERO ID: CRD42019122993.

Results: We identified nine randomized controlled trials and 16 observational studies. The fixed-effect model suggested a 0.99 probability that switch was the better strategy for increasing odds of a clinically meaningful improvement in ACR50 [odds ratio (OR): 1.35 (95% credible interval (CI): 0.96–1.81)]. The fixed-effect model also suggested that switch was associated with lower rates of withdrawal for any reasons [OR: 0.53 (95% CI: 0.40–0.68)]. The random-effects and hierarchical Bayesian models suggested additional uncertainty as they considered more variability than the fixed-effect model.

Discussion: Results suggest that switching to a drug with a different MOA is more effective and associated with lower rates of withdrawal than cycling to a different TNFi after failure of first-line TNFi. Further trials that directly compare cycling with switching are warranted to better assess comparative efficacy.

Correspondence to:
Alberto Migliore
Unit of Rheumatology,
Ospedale S. Pietro
Fatebenefratelli ISPOR
Italy, Via Cassia 600,
Rome, 00189, Italy
**migliore.alberto60@
gmail.com**

**Giuseppe Pompilio
Davide Integlia**
ISHEO Srl, Rome, Italy

**Joe Zhuo
Evo Alemao**
Worldwide Health
Economics & Outcomes
Research, Bristol Myers
Squibb, Princeton, NJ, USA

*At the time of the analysis

Plain language summary

Assessment of the effectiveness of different drug treatment strategies in patients with rheumatoid arthritis: an analysis of the published literature

- Rheumatoid arthritis (RA) is a chronic disease in which inflammation affects joints along with the entire body; this may cause significant pain, joint damage, physical disability, a decreased quality of life, and an increased risk of death.

- Tumor necrosis factor inhibitors (TNFis) are a common choice as first-line drugs to treat RA. Although they are effective in many patients, therapy with a TNFi is not successful within the first year of treatment in approximately one-third of patients due to either a lack of efficacy or safety issues.
- When TNFi therapy is unsuccessful, the options are to "cycle" to another TNFi or to "switch" to another drug with a different mechanism of action (MOA). Further studies are needed to help doctors decide the best treatment strategy for their patients when treatment with an initial TNFi fails.
- This study analyzed 25 published studies in which patients were either "cycled" to another TNFi or "switched" to a drug with a different MOA after unsuccessful treatment with an initial TNFi.
- The results showed that "switching" to a drug with a different MOA was a better treatment strategy than "cycling" to another TNFi; "switching" increased the chance of clinically meaningful improvement in disease status and lowered the chance of having to stop treatment for any reason.

Introduction

Rheumatoid arthritis (RA) is a chronic, debilitating disease characterized by persistent synovitis and systemic inflammation. When uncontrolled or untreated, RA can cause significant pain, joint erosion, systemic or extra-articular manifestations, functional disability, decreased quality of life, and an increased risk of death.[1] Conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), such as methotrexate, are considered the first line of treatment and are usually the first class of agents used to treat RA. For patients whose RA remains active even with csDMARDs therapy and for patients with poor prognostic factors, the consensus on treatment is to generally step-up to combination DMARD therapy or initiate the patient on a biologic agent.[2]

Currently the first choice of biologic therapy is typically a tumor necrosis factor inhibitor (TNFi), also known as an anti-TNF agent.[2] While TNFi therapy has been shown in randomized controlled trials (RCTs) to be effective at improving the signs and symptoms of RA, approximately 30–40% of the patients discontinue TNFi therapy within 1 year due either to failure of treatment efficacy or adverse events.[3,4] Patients with active disease despite TNFi therapy can subsequently cycle to either a different TNFi or switch to a biologic agent with an alternative mechanism of action (MOA). Currently, available data are limited for clinicians attempting to decide between these two strategies. Most clinical trials of second-line therapies for RA compare a specific second-line treatment with placebo. From existing trials,[5–10] we know that either cycling a patient on to a different TNFi or switching a patient to a drug with a different MOA can be beneficial. However, currently, it is not known which of these strategies is more effective.

This network meta-analysis (NMA) builds on a systematic literature review of RCTs and observational studies to investigate whether cycling or switching is more effective after TNFi failure in patients with RA.

Materials and methods

The protocol of our NMA was registered on PROSPERO international prospective register of systematic reviews (https://www.crd.york.ac.uk/PROSPERO/; ID = CRD42019122993). All relevant data are within the manuscript and its Supplemental files.

Literature search

The literature search was conducted in MEDLINE, Embase, and Cochrane Library in

June 2018. The search strategy can be found in Supplemental Appendix S1. RCTs and observational studies investigating second-line treatment for RA with TNFi (adalimumab, infliximab, etanercept, certolizumab, golimumab) and drugs with a different MOA (tofacitinib, rituximab, abatacept, tocilizumab, anakinra, baricitinib) in adult patients who had failed a previous treatment with TNFi were included in the present meta-analysis. The PICOS statement can be found in Supplemental Appendix S2 (Table S2-1).

Study selection

A systematic literature review was performed in two stages by two independent reviewers. Stage 1 encompassed review of titles and abstracts identified from the electronic search; stage 2 was based on a review of full-text articles of those deemed potentially relevant during stage 1. We excluded studies that did not provide information on the treatment type in at least one of the arms and those that did not report on the outcomes of interest. Non-English articles and non-original articles, including comments, editorials, case reports, and personal communications, were also excluded.

Data extraction

Data extraction was performed by two independent reviewers and any disagreements were resolved *via* consensus. The following information was extracted from the studies matching the inclusion criteria: the name of the author, year of publication, study design, demographic data of subjects, baseline clinical data of subjects, type of intervention, and numerical data on the outcomes of interest. The outcomes analyses were the mean reduction in health assessment questionnaire (HAQ) score, proportion of patients with a clinically meaningful HAQ reduction (defined as a reduction of at least 0.22), the proportion of patients with a 20% (ACR20), 50% (ACR50), or 70% (ACR70) response on the ACR score, the proportion of patients achieving a Disease Activity Score in 28 joints (DAS28) score below 2.6 or between 2.6 and 3.2, mean change in DAS28 score, rate of serious adverse events (SAEs), and number of withdrawals for any reason or due to adverse events or lack of treatment efficacy.

Quality assessment

We utilized the Cochrane Risk of Bias Tool[11] to evaluate the included RCTs[5–10,12–18] and the Newcastle–Ottawa Scale[19] to evaluate the observational studies.[20–35] The quality assessment was performed by two reviewers independently, and any disagreement was resolved *via* consensus.

Classification of treatment

There is currently no consensus on the optimal next treatment after TNFi failure.[36] The two main options are to cycle patients on to a different TNFi, or to switch patients on to a drug with a different MOA. Therefore, in this study, treatments were classified as either cycle or switch. Interventions in which patients moved to a different TNFi were classed as cycle. Interventions in which patients moved to a treatment with a different MOA were classed as switch. Interventions in which patients stayed on the same drug were excluded. Placebo was classed as placebo [Supplemental Appendix S2 (Table S2-2)].

Network meta-analysis

For each endpoint, we described the geometry of the network by drawing network diagrams [Supplemental Appendix S3 (Figures S3-1 to S3-13)] that illustrate the different treatments and the number of trials and/or observational studies used in each comparison. Separate diagrams were necessary because the different studies often reported only a selection of the outcomes of interest.

Statistical model

Because this analysis takes data from studies covering a range of populations and study designs, the decision was made to develop three models to conduct the network meta-analyses: a fixed-effect model, a random-effects model, and a hierarchical Bayesian model. The first two models compared cycle treatments *versus* switch treatments, and the hierarchical model compared the individual drugs against each other while clustered into one of the two treatment classes. A hierarchical model in turn allowed for the individual drugs within each class to be modeled separately in sub-models that are then clustered into the treatment classes.[37] Using these three types of models enabled us to compare the primary results and choose based on the deviance information criterion (DIC).

For each endpoint, we performed three NMAs. These models were built in R v3.5.1 and WinBUGS v1.4.3. The three NMAs were:

1. Fixed-effect NMA of treatment class (i.e. cycle *versus* switch *versus* placebo)
2. Random-effects NMA of treatment class (i.e. cycle *versus* switch *versus* placebo)
3. Hierarchical Bayesian NMA of treatment drug, treating each drug as clustered within its treatment class [Supplemental Appendix S2 (Table S2-2)]

We chose primary results between fixed-effect and random-effects network meta-analyses by considering the DIC: where the fixed-effect meta-analysis had a random-effect DIC +5 or less, we reported the fixed-effect results as per standard practice.[38]

For continuous outcomes, we used linear regression models. For binary analyses, we used logistic regression models.

Assessing consistency and heterogeneity

There were no RCTs of a cycle drug *versus* a switch drug, so we could not assess the consistency of analyses of RCTs.

For the full analysis, including both RCTs and observational studies, consistency was assessed where possible by fitting models with consistency factors. For the consistency model, a DIC 5 or more points below the DIC of the standard fixed-effects model was regarded as an important lack of consistency. Heterogeneity among the comparative studies was assessed using the I^2 statistic.

Choice of priors for each model

For the binary model, we used a diffuse prior for trial baselines on the logit scale $\mu_i Normal(0,100^2)$ and weakly informative priors for treatment effects $d_k Normal(0,3.3^2)$. In the random-effects model and hierarchical model, we used a diffuse prior for the between-trial precision $\tau Gamma(0.001,0.001)$ (the between-trial standard deviation $sd = \frac{1}{\sqrt{\tau}}$). In the hierarchical model, we used a weakly informative prior for the treatments within a specific treatment class $D.d_k Normal(0,3.3^2)$.

For the continuous model, we used a diffuse prior for trial baselines $\mu_i Normal(0,100^2)$ and diffuse priors for treatment effects $d_k Normal(0,100^2)$. In the random-effects model and hierarchical model, we used a diffuse prior for the between-trial precision $\tau Gamma(0.01,0.01)$ (the between-trial standard deviation $sd = \frac{1}{\sqrt{\tau}}$). In the hierarchical model, we used a weakly informative prior for the treatments within a specific treatment class $D.d_k Normal(0,10^2)$.

For each model, we ran 20,000 iterations, discarding the first 10,000 iterations as burn-in. This was an excessive number of iterations, used because estimates of all considered factors had reached convergence (Rhat == 1) and computer time was acceptable (<30 s per model).

Presentation of results

For each NMA, we estimated pair-wise mean differences (continuous endpoints) or odds ratios (ORs; binary endpoints) and used these to present results, accompanied with appropriate 95% credible intervals (CIs). We also present the probability that a certain treatment class is the best in a certain measure. This means that a higher probability indicates that the treatment class is likely better at achieving a positive outcome, such as achieving an improvement in HAQ score, or better at reducing a negative outcome, such as a SAE or a withdrawal.

We additionally estimated the surface under the cumulative ranking (SUCRA) and average treatment rankings to explore potential treatment hierarchy. SUCRA is a numerical presentation of the overall probability that a drug will occupy at least one of the top ranks, with possible values ranging from 0% to 100%. Using SUCRA to rank treatments must be done with caution, however, as these rankings can arise from evidence that has low certainty. A set of SUCRAs can be the same whether they come from high-quality studies or low-quality studies; the method itself does not differentiate. Therefore, SUCRA results are only as robust as the data on which they are based.[39]

Average treatment rankings indicate the mean ranking of the treatment across the iterations of the model. Combining the average ranking with the SUCRA rating provides an indication of the uncertainty in the results: if a treatment has a high SUCRA probability but a low average ranking, or *vice versa*, these divergent results may reflect a lack of certainty and be a result of a treatment
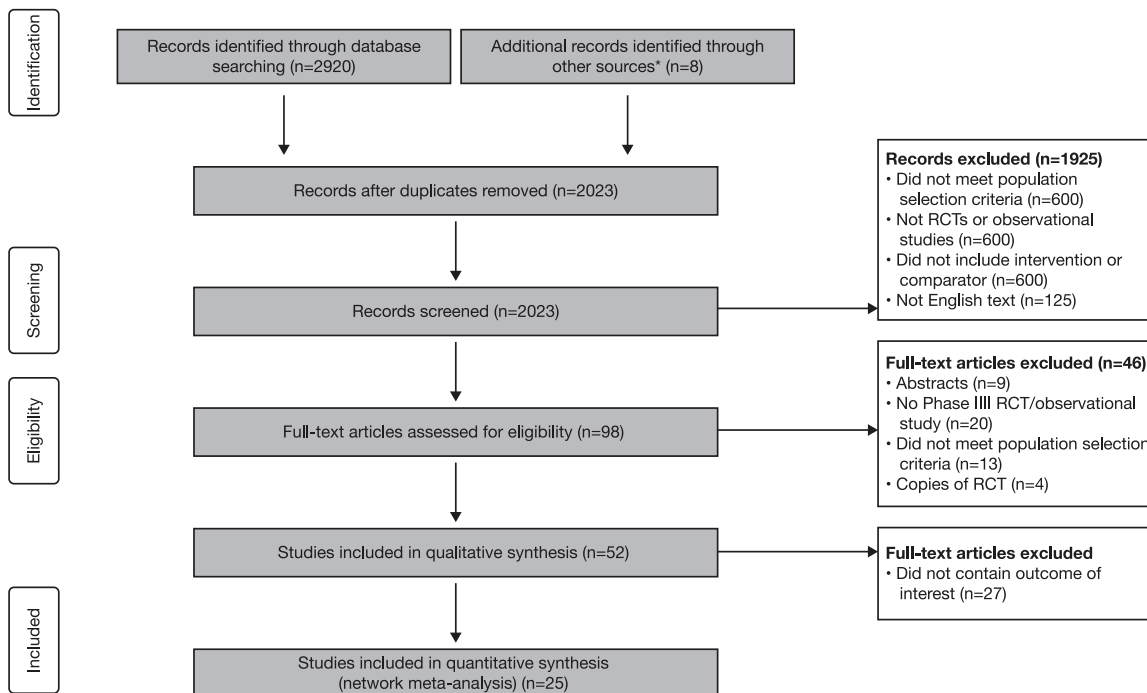
Figure 1. PRISMA flow diagram.
*Free searches using Google Scholar and PubMed.
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial.

having a high probability of being in the top ranks or the bottom ranks because of underlying uncertainty.

Results

Literature search
Of the 2928 articles identified during the literature search, 2830 were excluded and the full text of the remaining 98 articles was reviewed for eligibility. After the full texts were reviewed, 46 articles were excluded after meeting the exclusion criteria and 27 articles were excluded as they did not contain the outcome of interest. We have included in the NMAs nine RCTs reported in 13 articles (four for cycling and five for switching) and 16 observational studies (eight for cycling and eight for switching; Figure 1). As the studies on baricitinib did not meet the criteria reported in PICOS statement, it was not included in the NMA.

Study characteristics
A total of 25 studies were included in the NMA, the baseline characteristics of which are summarized in Supplemental Appendix S4 (Table S4-1).

The baseline characteristics varied across the studies, with patient numbers ranging from small studies of 15 patients up to 1683 patients. The proportion of female participants in the studies was at least 71.1% and the mean age of participants ranged from 45.1 to 59.0 years. Most patients had been living with RA for a long time, with the mean disease duration ranging from 6.8 to 16.6 years. Baseline clinical measures were generally more severe in the RCTs than in the observational studies. Examples of this included the mean tender joint count and swollen joint count, which ranged from 22.8 to 33.9 and 12.8 to 23.4, respectively, in the RCTs, and 8.2 to 17.8 and 5.9 to 14.3, respectively, in the observational studies.

Assessment of risk of bias
The results of the risk of bias assessments for both the RCTs and the observational studies are presented in Supplemental Appendix S4 (Figures S4-1 and S4-2). While many of the RCTs were constructed to minimize bias, four did have aspects that introduced bias into the results.[9,12,14,15] The bias in these studies can mainly be attributed to performance bias from the lack of blinding of study participants and outcome assessors, and
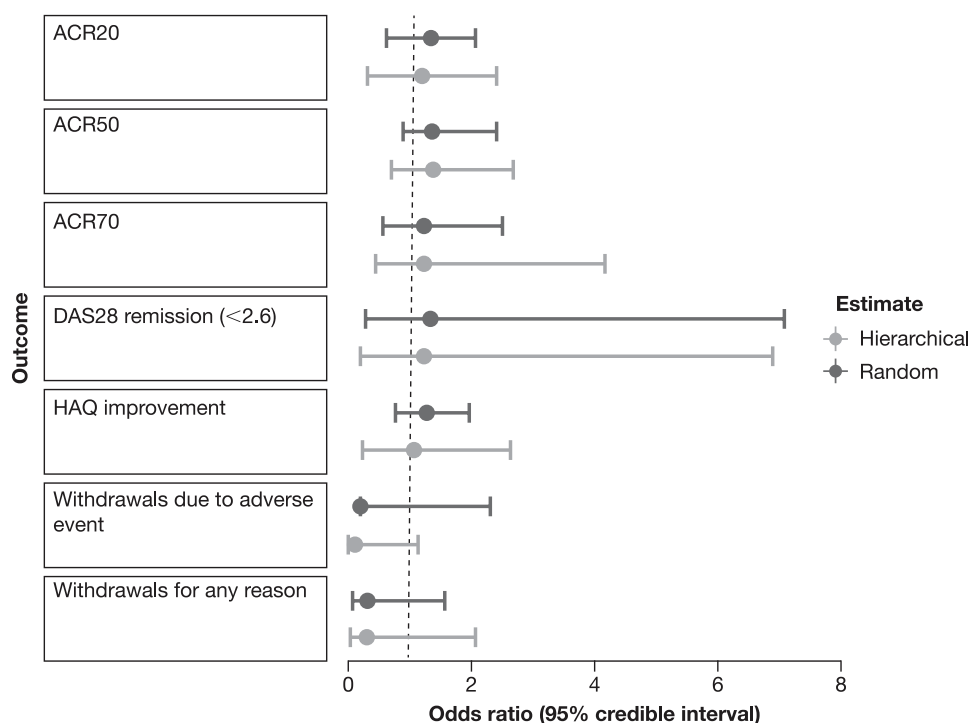
Figure 2. Forest plot showing Bayesian hierarchical and random-effects odds ratio switch *versus* cycle estimate for the main outcomes.
ACR20/50/70, American College of Rheumatology criteria measuring a 20/50/70% improvement on a scale of 28 intervals; DAS28, Disease Activity Score in 28 joints; HAQ, Health Assessment Questionnaire.

selection bias due to the open-label nature of the studies.

The bias introduced in the observational studies was mainly attributed to the lack of comparability of the cohorts involved. For 11 of the studies,[20–25,27,30,33–35] there was no attempt to improve the comparability of the cohorts as the only differentiation used was previous treatment received.

Heterogeneity
There was no evidence of significant heterogeneity when exploring the HAQ improvement, ACR20 and DAS28 remission outcomes in the studies that compared either a switch or a cycle approach with placebo. However, this was not the case when comparing switch with cycle. For the HAQ improvement outcome, the I^2 value was 51.69% and for the DAS28 outcome the I^2 value was 97.05% [Supplemental Appendix S4 (Figure S4-3)]. This heterogeneity is likely due to only two studies being used, the smaller of which[28] showed a statistically significant improvement based on log OR for switch compared with cycle

and the other, larger study[32] demonstrating much less of an improvement.

Efficacy outcomes
The results of key outcomes showed that the switch approach had a higher probability of resulting in a better outcome than the cycle approach (Figure 2), as highlighted in Figure 3 for the main outcomes (ACR20, ACR50, ACR70, DAS28 remission, HAQ improvement, withdrawals for any reason, and withdrawals due to adverse events). Results for the key outcomes of interest are presented in the main manuscript and Supplemental Appendix S5; other outcomes (DAS28 low, DAS28 erythrocyte sedimentation rate change, DAS28 C-reactive protein change, HAQ reduction, withdrawals due to lack of efficacy, SAEs) are reported in Supplemental Appendix S6.

Figure 2 shows the OR switch *versus* cycle for the above-mentioned main outcomes for both random and hierarchical Bayesian estimate. For the positive outcomes (ACR20, ACR50, ACR70,
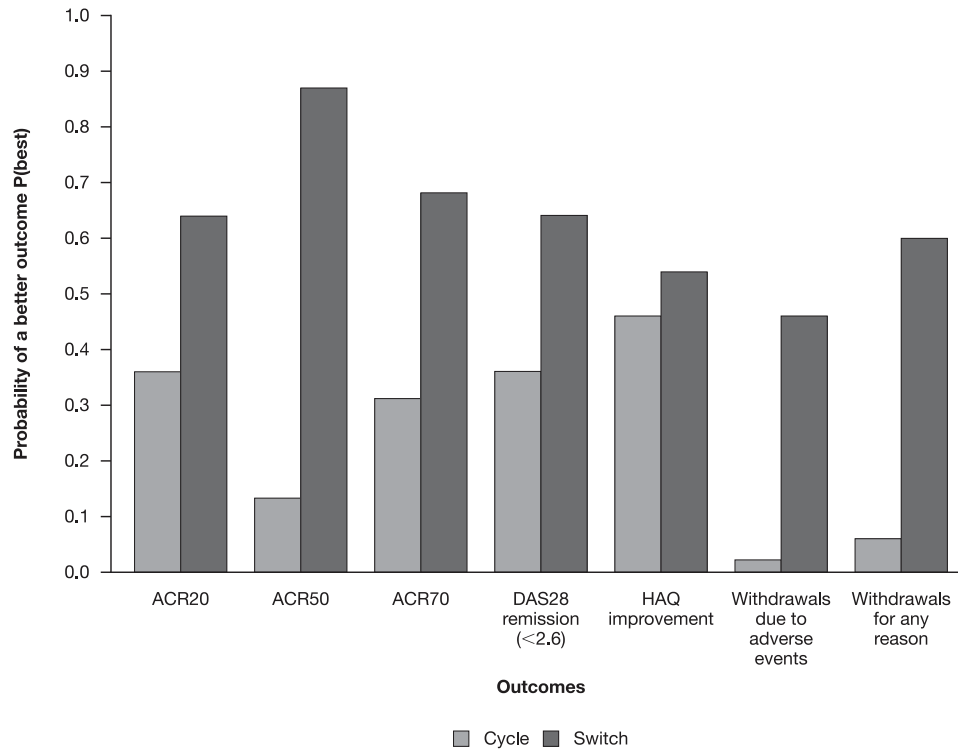
Figure 3. Bayesian hierarchical model of the probability of a better outcome from the cycle or switch approach for the main outcomes (ACR20, ACR50, ACR70, DAS28 remission, HAQ improvement, withdrawals due to adverse events and withdrawals for any reason).
ACR20/50/70, American College of Rheumatology criteria measuring a 20/50/70% improvement on a scale of 28 intervals; DAS28, Disease Activity Score in 28 joints; HAQ, Health Assessment Questionnaire.

DAS28 remission, HAQ improvement) the OR was in favor of the switch *versus* cycle approach. ORs for negative outcomes (withdrawals for any reason and withdrawals due to adverse event) were also in favor of a switch approach. It is worth noting that the estimates are all to the left of the point of indifference. As far as the Bayesian analysis is concerned, it should be considered that the criteria for interpreting CIs used in frequency analysis are much less relevant in Bayesian analysis. Therefore, even if the CIs cross the line of indifference, we consider the result equally valid in probabilistic terms.

ACR20
ACR20 was reported by seven studies: six RCTs and one observational study [Supplemental Appendix (Figure S3-1)]. The Bayesian hierarchical NMA indicated a 64% probability that switch is better than cycle for ACR20 [Supplemental Appendix S5 (Table S5-1)], with fixed- and random-effects class-based analyses showing much higher probabilities. It was unclear

which specific switch drug was best [Supplemental Appendix (Table S5-2)]. Results based only on RCTs differed, with the fixed-effect model showing a 76% probability that switch is better than cycle, and the hierarchical model showing a 61% probability that cycle is better than switch. The random-effects model was split, giving a 51% probability that switch is better than cycle [Supplemental Appendix S5 (Table S5-3)].

There was no evidence of inconsistency (DIC of fixed-effect inconsistency model was 100 *versus* 101 for the consistency model).

ACR50
ACR50 was reported by eight studies: six RCTs and two observational studies [Supplemental Appendix (Figure S3-2)]. The Bayesian hierarchical NMA gave an 87% probability that switch is better than cycle for ACR50 (Table 1), with fixed- and random-effects class-based analyses showing much higher probabilities. The fixed-effect model, for example, suggested a 99%

Table 1. Results of network meta-analysis of ACR50 based on all studies reporting ACR50.

Statistic	Fixed-effect estimate	Lower bound	Upper bound	Random-effect estimate	Lower bound	Upper bound	Hierarchical model estimate	Lower bound	Upper bound
DIC	76			77			98		
OR cycle *versus* placebo	3.86	2.53	6.2	3.82	1.99	6.55	3.84	1.87	7.39
OR switch *versus* placebo	5.15	3.68	7.52	5.27	3.31	8.29	5.33	3.26	8.82
OR switch *versus* cycle	1.35	0.96	1.81	1.38	0.88	2.41	1.39	0.69	2.69
P(placebo best)	0			0			0		
P(cycle best)	0.04			0.06			0.13		
P(switch best)	0.96			0.94			0.87		
SUCRA(placebo)	0			0			0		
SUCRA(cycle)	0.52			0.53			0.56		
SUCRA(switch)	0.98			0.97			0.94		
Average rank (placebo)	3			3			3		
Average rank (cycle)	1.96			1.94			1.87		
Average rank (switch)	1.04			1.06			1.13		

ACR50, American College of Rheumatology criteria measuring a 50% improvement on a scale of 28 intervals; DIC, deviance information criterion; OR, odds ratio; SUCRA, surface under the cumulative ranking curve.

probability that switch is a better strategy for increasing the odds of a clinically meaningful improvement in ACR50 [OR: 1.35 (95% CI: 0.96–1.81)]. It was unclear which specific switch drug was best [Supplemental Appendix S5 (Table S5-4)]. Results based only on RCTs were very similar with only a slight reduction in probability that switch is the best [Supplemental Appendix S5 (Table S5-5)].

There was no evidence of inconsistency (DIC of fixed-effect inconsistency model was 76 *versus* 77 for the consistency model).

ACR70
ACR70 was reported in six studies: five RCTs and one observational study [Supplemental Appendix S3 (Figure S3-3)].

The Bayesian hierarchical NMA gave a 68% probability that switch is better than cycle for ACR70 [Supplemental Appendix S5 (Table S5-6)], with fixed and random-effects class-based

analysis giving higher probabilities. It was unclear which specific switch drug was best [Supplemental Appendix S5 (Table S5-7)]. Results based only on RCTs were similar [Supplemental Appendix S5 (Table S5-8)].

There was no evidence of inconsistency (DIC of fixed-effect inconsistency model was 65 *versus* 66 for the consistency model).

DAS28 remission
DAS28 remission proportion was reported in five RCTs [Supplemental Appendix S3 (Figure S3-4)].

The Bayesian hierarchical NMA gave a 64% probability that switch is better than cycle for DAS28 remission (Table 2), with fixed- and random-effects class-based analyses giving higher probabilities. It was unclear which specific switch drug was best [Supplemental Appendix S5 (Table S5-9)].

A consistency model was not possible.

Table 2. Results of network meta-analysis of DAS28 remission based on all studies reporting DAS28 remission.

Statistic	Fixed-effect estimate	Lower bound	Upper bound	Random-effect estimate	Lower bound	Upper bound	Hierarchical model estimate	Lower bound	Upper bound
DIC	103			82			90		
OR cycle *versus* placebo	7.59	3.7	17.57	8.3	1.39	51.52	9.02	1.45	52.5
OR switch *versus* placebo	10.29	5.26	23.99	11.26	2.7	43.66	11.21	2.39	52.59
OR switch *versus* cycle	1.37	1.04	1.83	1.33	0.28	7.09	1.23	0.2	6.9
P(placebo best)	0			0			0		
P(cycle best)	0.01			0.3			0.36		
P(switch best)	0.99			0.7			0.64		
SUCRA(placebo)	0			0.01			0.01		
SUCRA(cycle)	0.51			0.64			0.67		
SUCRA(switch)	0.99			0.85			0.82		
Average rank (placebo)	3			2.98			2.98		
Average rank (cycle)	1.99			1.71			1.65		
Average rank (switch)	1.01			1.31			1.37		

DAS28, Disease Activity Score in 28 joints; DIC, deviance information criterion; OR, odds ratio; SUCRA, surface under the cumulative ranking curve.

HAQ improvement

HAQ improvement was reported in eight studies: six RCTs and two observational studies [Supplemental Appendix S3 (Figure S3-4)].

The balance of evidence suggested that switch is better than cycle for HAQ improvement, with the random-effects model giving this an 88% probability [Supplemental Appendix S5 (Table S5-10)]. This probability was lower than in the class-based analysis due to the inclusion of treatments such as anakinra (switch) and adalimumab (cycle) for which no studies reported HAQ improvement, and, therefore, there is uncertainty around their impact on HAQ improvement. This uncertainty also impacted determining which drug was best; infliximab (cycle) was the most likely with P(best) = 0.22 and SUCRA = 0.67, but tocilizumab (switch) and rituximab (switch) also showed promise with P(best) = 0.14 and SUCRA = 0.67 and P(best) = 0.09 and SUCRA = 0.69 respectively (Supplemental Appendix S5 [Table S5-11]).

Results based on RCTs were similar [Supplemental Appendix S5 (Table S5-12)].

There was no evidence of inconsistency (DIC of fixed-effect inconsistency model was 105 *versus* 104 for the consistency model).

Safety outcomes

Withdrawals for any reason. Withdrawals for any reason were reported by eight studies: five RCTs and three observational studies [Supplemental Appendix S3 (Figure S3-6)]. The fixed-effect model suggested that switch was associated with lower rates of withdrawal for any reasons [OR: 0.53 (95% CI: 0.40–0.68); Supplemental Appendix S5 (Table S5-13)].

The Bayesian hierarchical NMA gave a 60% probability that switch leads to fewer withdrawals for any reason than cycle or placebo [Supplemental Appendix S5 (Figure S5-13)], with fixed- and random-effects class-based analyses showing similar probabilities. It is unclear which specific treatments led to the highest or lowest number of withdrawals, with cycle and switch treatments sharing similar probabilities compared with other treatments within their classes [Supplemental

Table 3. Results of network meta-analysis of withdrawals due to adverse events based on all studies reporting withdrawals due to adverse events.

Statistic	Fixed-effects estimate	Lower bound	Upper bound	Random-effects estimate	Lower bound	Upper bound	Hierarchical model estimate	Lower bound	Upper bound
DIC	63			64			75		
OR cycle *versus* placebo	12.17	2.92	78.86	9.84	0.30	64.57	8.03	0.83	55.04
OR switch *versus* placebo	1.09	0.62	1.80	1.02	1.02	2.35	1.03	0.43	2.62
OR switch *versus* cycle	0.09	0.01	0.35	0.2	0.19	2.30	0.13	0.02	1.15
P(placebo best)	0.62			0.50			0.52		
P(cycle best)	0			0.04			0.02		
P(switch best)	0.38			0.46			0.46		
SUCRA(placebo)	0.81			0.73			0.75		
SUCRA(cycle)	0.00			0.06			0.03		
SUCRA(switch)	0.69			0.71			0.72		
Average rank (placebo)	1.38			1.54			1.50		
Average rank (cycle)	3.00			2.89			2.94		
Average rank (switch)	1.62			1.57			1.56		

AE, adverse event; DIC, deviance information criterion; OR, odds ratio; SUCRA, surface under the cumulative ranking curve.

Appendix S5 (Table S5-14)]. Results based on RCTs only were similar [Supplemental Appendix S5 (Table S5-15)].

There was evidence of inconsistency (DIC of fixed-effect inconsistency model was 112 *versus* 83 for the consistency model). Two observational studies directly compared cycle with switch: of the 801 patients receiving an unspecified TNFi, 391 withdrew due to lack of efficacy, while of the 355 patients receiving rituximab, 111 withdrew, resulting in an OR of 0.47 for switch *versus* cycle. The indirect evidence (the results of the RCTs) gave an OR for switch *versus* cycle of 0.86 [Supplemental Appendix S5 (Table S5-15)]. Given that the direct evidence is based on a large number of patients in the two observational studies, the difference in ORs between the direct evidence from the observational studies and the indirect evidence from the RCT likely explains the apparent inconsistency. Because the direct evidence is based on a comparison of one switch

treatment, rituximab, against unspecified cycle treatments, while the RCTs include a wider range of switch treatments, it is likely the RCT data are more robust.

Withdrawals due to adverse events. Withdrawals due to adverse events were reported by eight studies: six RCTs and two observational studies [Supplemental Appendix S3 (Figure S3-7)].

The hierarchical Bayesian NMA of all studies gave a 46% probability that switch causes fewer withdrawals due to adverse events than cycle (Table 3). Results for the switch fixed- and random-effects models were also higher than those for cycle [Supplemental Appendix S5 (Table S5-14)].

Switch treatments showed the highest probabilities for leading to the lowest number of withdrawals, with P(best) = 0.14 and P(best) = 0.17 for abatacept and anakinra, respectively [Supplemental Appendix S5 (Table S5-16)].

There was no evidence of inconsistency (DIC of fixed-effect inconsistency model was 63.0 *versus* 64 for the consistency model).

Limitations
Limitations of this study include the limited number of head-to-head studies (cycle *versus* switch) included in it, potential bias introduced from the lack of blinding in four of the included RCTs and from the inability to ensure comparability between cohorts in 11 of the observational studies.

Discussion
In recent years, alongside the common clinical practice of cycling among biological therapies in patients with RA who have an inadequate response to TNFi drugs,[40,41] the therapeutic strategy of switching to biologic agents with different MOAs has also become a common practice.[42,43]

The objective of this study was to compare the efficacy and safety of cycling patients to another TNFi *versus* switching patients to a drug with a different MOA in adult patients with RA who had failed previous treatment with TNFi. Data from RCTs and observational studies were adjusted for baseline characteristics and were used to perform a Bayesian NMA; consistency and heterogeneity were also evaluated.

Across all considered efficacy outcomes (ACR20, ACR50, ACR70, DAS28 score below 2.6 or between 2.6 and 3.2, mean change in DAS28 score, mean and reduction in HAQ score) results consistently showed a higher probability that switching is a better strategy than cycling. Further, in the hierarchical model switching was also associated with lower negative outcomes, such as number of withdrawals for any reasons or due to adverse events, and lack of efficacy. The switch strategy showed uncertain results for SAEs because the hierarchical analysis with all the studies and the analysis with only the RCTs went slightly in the opposite direction, even if the results were very close. In general, the probability that switching was better than cycling was in the range of 70–90% rather than being certain, while it was unclear which drug was most effective.

Our results suggest that switching to a drug with a different MOA is more effective and associated with lower rates of withdrawals than cycling to a different TNFi after failure of first-line TNFi

therapy. Further trials to compare cycling with switching directly are warranted to better assess the comparative efficacy of these two treatment strategies.

Conflict of interest statement
AM acted as a consultant for Bristol Myers Squibb, Fidia Pharma, IBSA, Merck, Pfizer, Roche, Sanofi-Aventis and UCB.

GP is an employee of ISHEO SRL.

DI received grant/research support from AbbVie, Angelini, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly, Fidia Pharma, Merck Serono and Pierre Fabre, and is an employee of ISHEO SRL.

JZ is an employee of and shareholder in Bristol Myers Squibb.

EA was an employee of and shareholder in Bristol Myers Squibb at the time of the analysis.

ORCID iDs
Alberto Migliore (iD) https://orcid.org/0000-0002-7631-890X

Giuseppe Pompilio (iD) https://orcid.org/0000-0002-0971-4558

Supplemental material
Supplemental material for this article is available online.

References
1. Scott DL, Wolfe F and Huizinga TWJ. Rheumatoid arthritis. *Lancet* 2010; 376: 1094–1108.

2. Smolen JS, Landewe RBM, Bijlsma JWJ, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and

biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020; 79: 685–699.

3. Favalli EG, Raimondo MG, Becciolini A, *et al.* The management of first-line biologic therapy failures in rheumatoid arthritis: current practice and future perspectives. *Autoimmun Rev* 2017; 16: 1185–1195.

4. Soliman MM, Ashcroft DM, Watson KD, *et al.* Impact of concomitant use of DMARDs on the persistence with anti-TNF therapies in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. *Ann Rheum Dis* 2011; 70: 583–589.

5. Emery P, Keystone E, Tony HP, *et al.* IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. *Ann Rheum Dis* 2008; 67: 1516–1523.

6. Smolen JS, Kay J, Doyle MK, *et al.* Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor α inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial. *Lancet* 2009; 374: 210–221.

7. Keystone E, Emery P, Peterfy CG, *et al.* Rituximab inhibits structural joint damage in patients with rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitor therapies. *Ann Rheum Dis* 2009; 68: 216–221.

8. Genovese MC, Becker J-C, Schiff M, *et al.* Abatacept for rheumatoid arthritis refractory to tumor necrosis factor α inhibition. *N Engl J Med* 2005; 353: 1114–1123.

9. Schiff MH, von Kempis J, Goldblum R, *et al.* Rheumatoid arthritis secondary non-responders to TNF can attain an efficacious and safe response by switching to certolizumab pegol: a phase IV, randomised, multicentre, double-blind, 12-week study, followed by a 12-week open-label phase. *Ann Rheum Dis* 2014; 73: 2174–2177.

10. Burmester GR, Blanco R, Charles-Schoeman C, *et al.* Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. *Lancet* 2013; 381: 451–460.

11. Sterne JAC, Savovic J, Page MJ, *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; 366: l4898.

12. Bykerk VP, Ostor AJ, Alvaro-Gracia J, *et al.* Tocilizumab in patients with active rheumatoid arthritis and inadequate responses to DMARDs and/or TNF inhibitors: a large, open-label study close to clinical practice. *Ann Rheum Dis* 2012; 71: 1950–1954.

13. Cohen SB, Emery P, Greenwald MW, *et al.* Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum* 2006; 54: 2793–2806.

14. Fleischmann R, Goldman JA, Leirisalo-Repo M, *et al.* Infliximab efficacy in rheumatoid arthritis after an inadequate response to etanercept or adalimumab: results of a target-driven active switch study. *Curr Med Res Opin* 2014; 30: 2139–2149.

15. Furst DE, Gaylis N, Bray V, *et al.* Open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept: the opposite study. *Ann Rheum Dis* 2007; 66: 893–899.

16. Keystone E, Burmester GR, Furie R, *et al.* Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Rheum* 2008; 59: 785–793.

17. Strand V, Burmester GR, Ogale S, *et al.* Improvements in health-related quality of life after treatment with tocilizumab in patients with rheumatoid arthritis refractory to tumour necrosis factor inhibitors: results from the 24-week randomized controlled RADIATE study. *Rheumatology (Oxford)* 2012; 51: 1860–1869.

18. Westhovens R, Cole JC, Li T, *et al.* Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. *Rheumatology (Oxford)* 2006; 45: 1238–1246.

19. Wells GA, Shea B, O'Connell D, *et al.* The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2016. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (acessed 25 April 2016).

20. Backhaus M, Kaufmann J, Richter C, *et al.* Comparison of tocilizumab and tumour necrosis factor inhibitors in rheumatoid arthritis: a retrospective analysis of 1603 patients managed in routine clinical practice. *Clin Rheumatol* 2015; 34: 673–681.

21. Bingham CO III, Ince A, Haraoui B, *et al.* Effectiveness and safety of etanercept in subjects

with RA who have failed infliximab therapy: 16-week, open-label, observational study. *Curr Med Res Opin* 2009; 25: 1131–1142.

22. Blom M, Kievit W, Donders ART, *et al.* Effectiveness of a third tumor necrosis factor-α-blocking agent compared with rituximab after failure of 2 TNF-blocking agents in rheumatoid arthritis. *J Rheumatol* 2011; 38: 2355–2361.

23. Bombardieri S, Ruiz AA, Fardellone P, *et al.* Effectiveness of adalimumab for rheumatoid arthritis in patients with a history of TNF-antagonist therapy in clinical practice. *Rheumatology (Oxford)* 2007; 46: 1191–1199.

24. Chatzidionysiou K, Askling J, Eriksson J, *et al.*; ARTIS group. Effectiveness of TNF inhibitor switch in RA: results from the national Swedish register. *Ann Rheum Dis* 2015; 74: 890–896.

25. Emery P, Gottenberg JE, Rubbert-Roth A, *et al.* Rituximab versus an alternative TNF inhibitor in patients with rheumatoid arthritis who failed to respond to a single previous TNF inhibitor: SWITCH-RA, a global, observational, comparative effectiveness study. *Ann Rheum Dis* 2015; 74: 979–984.

26. Favalli G, Arreghini M, Arnoldi C, *et al.* Anti-tumor necrosis factor a switching in rheumatoid arthritis and juvenile chronic arthritis. *Arthritis Rheum* 2004; 51: 301–302.

27. Gomez-Reino JJ, Maneiro JR, Ruiz J, *et al.* Comparative effectiveness of switching to alternative tumour necrosis factor (TNF) antagonists versus switching to rituximab in patients with rheumatoid arthritis who failed previous TNF antagonists: the MIRAR Study. *Ann Rheum Dis* 2012; 71: 1861–1864.

28. Harrold LR, Reed GW, Magner R, *et al.* Comparative effectiveness and safety of rituximab versus subsequent anti-tumor necrosis factor therapy in patients with rheumatoid arthritis with prior exposure to anti-tumor necrosis factor therapies in the United States Corrona registry. *Arthritis Res Ther* 2015; 17: 256.

29. Harrold LR, Reed GW, Solomon DH, *et al.* Comparative effectiveness of abatacept versus tocilizumab in rheumatoid arthritis patients with prior TNFi exposure in the US Corrona registry. *Arthritis Res Ther* 2016; 18: 280.

30. Kekow J, Mueller-Ladner U and Schulze-Koops H. Rituximab is more effective than second anti-TNF therapy in rheumatoid arthritis patients and previous TNFα blocker failure. *Biologics* 2012; 6: 191–199.

31. Nikas SN, Voulgari PV, Alamanos Y, *et al.* Efficacy and safety of switching from infliximab to adalimumab: a comparative controlled study. *Ann Rheum Dis* 2006; 65: 257–260.

32. Soliman MM, Hyrich KL, Lunt M, *et al.* Rituximab or a second anti-tumor necrosis factor therapy for rheumatoid arthritis patients who have failed their first anti-tumor necrosis factor therapy? Comparative analysis from the British Society for Rheumatology Biologics Register. *Arthritis Care Res (Hoboken)* 2012; 64: 1108–1115.

33. van der Bijl AE, Breedveld FC, Antoni CE, *et al.* An open-label pilot study of the effectiveness of adalimumab in patients with rheumatoid arthritis and previous infliximab treatment: relationship to reasons for failure and anti-infliximab antibody status. *Clin Rheumatol* 2008; 27: 1021–1028.

34. Virkki LM, Valleala H, Takakubo Y, *et al.* Outcomes of switching anti-TNF drugs in rheumatoid arthritis--a study based on observational data from the Finnish Register of Biological Treatment (ROB-FIN). *Clin Rheumatol* 2011; 30: 1447–1454.

35. Wick MC, Ernestam S, Lindblad S, *et al.* Adalimumab (Humira) restores clinical response in patients with secondary loss of efficacy from infliximab (Remicade) or etanercept (Enbrel): results from the STURE registry at Karolinska University Hospital. *Scand J Rheumatol* 2005; 34: 353–358.

36. Gottenberg J-E, Brocq O, Perdriger A, *et al.* Non–TNF-targeted biologic vs a second anti-TNF drug to treat rheumatoid arthritis in patients with insufficient response to a first anti-TNF drug: a randomized clinical trial. *JAMA* 2016; 316: 1172–1180.

37. Gelman A, Carlin JB, Stern HS, *et al. Bayesian data analysis.* 2nd ed. Boca Raton, FL: Chapman and Hall/CRC, 2004.

38. Spiegelhalter DJ, Best NG, Carlin BP, *et al.* Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol* 2002; 64: 583–639.

39. Mbuagbaw L, Rochwerg B, Jaeschke R, *et al.* Approaches to interpreting and choosing the best treatments in network meta-analyses. *Syst Rev* 2017; 6: 79.

40. Desai RJ, Solomon DH, Jin Y, *et al.* Temporal trends in use of biologic DMARDs for rheumatoid arthritis in the United States: a cohort study of publicly and privately insured patients. *J Manag Care Spec Pharm* 2017; 23: 809–814.

41. Tvete IF, Natvig B, Gasemyr J, *et al.* Comparing effects of biologic agents in treating patients

with rheumatoid arthritis: a multiple treatment comparison regression analysis. *PLoS One* 2015; 10: e0137258.

42. Jin Y, Desai RJ, Liu J, *et al.* Factors associated with initial or subsequent choice of biologic disease-modifying antirheumatic drugs for treatment of rheumatoid arthritis. *Arthritis Res Ther* 2017; 19: 159.

43. Reynolds A, Koenig AS, Bananis E, *et al.* When is switching warranted among biologic therapies in rheumatoid arthritis? *Expert Rev Pharmacoecon Outcomes Res* 2012; 12: 319–333.