Study	Incomplete outcome data	Incomplete outcome data	Free of Selective Reporting	Free of Selective Reporting	Other biases- major baseline imbalance	Other biases- major baseline imbalance
	Judgement	Reason	Judgement	Reason	Judgement	Reason
Abe 2006	Low Risk	"Out of 151 patients enrolled in the DBT, 147 received at least one infusion of study drugs (47, 49, and 51 patients in the placebo, 3 mg/kg, and 10 mg/kg groups, respectively). Five patients receiving the placebo discontinued treatment, including 3 due to lack of efficacy, one due to an adverse event, and one due to a protocol violation. Five patients receiving infliximab discontinued treatment due to adverse events."	Unclear	Not enough information provided to assess selective reporting	Low Risk	Baseline demographics comparable among the 3 groups
Bae 2013	High Risk	Patient reported outcomes data were analyzed from the mITT population, consisting of all patients who received at least one dose of study drug (ETN or DMARD) in combination with MTX and submitted at least one postbaseline assessment. Missing data were imputed using the last observation carried forward (LOCF) method for HAQ scores."	Unclear	Not enough information provided to assess selective reporting	Low Risk	"Demographic and baseline disease characteristics were not significantly different between the two treatment groups"
Chen 2009	High Risk	"The efficacy analysis was performed on an “intentto- treat (ITT)” population, which was defined as all patients with baseline data and at least one posttreatment evaluation. The last observation carried forward (LOCF) method was used to substitute for missing data."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Choy 2012	High Risk	Frequency of withdrawal "much higher in PL plus MTX group than the CZP plus MTX group."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Cohen 2002	High Risk	Full intention-to-treat analysis was performed on analysis of results. However, for the patients consenting to the 24 week trial, there was a 21.0% withdrawal rate. However, if the patients from the original 12 week trial who refused to consent for the full 24 week study extension were included, only 45.3% of the original study participants followed up for the full 24 weeks.	Unclear	Not enough information provided to assess selective reporting	Low Risk	Balanced baseline characteristics
Cohen 2004	Low Risk	Intention-to-treat analysis performed	Unclear	Not enough information provided to assess selective reporting	Low Risk	Balanced baseline characteristics
Combe 2006	Unclear	No data presented	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Combe 2009	High Risk	"Efficacy analyses were based on a modified intent-to-treat population, including patients who received any test article and provided efficacy data at baseline and at any subsequent visit… The last-observation-carried-forward (LOCF) approach was used to account for missing data points. LOCF imputation was also applied to patients who discontinued for unsatisfactory response."	Low Risk	Not enough information provided to assess selective reporting	High Risk	"There were no significant differences among the groups in the baseline characteristics with the exception of the percentage of patients receiving previous corticosteroids and the mean number of previous DMARD"
Conaghan 2013	Low Risk	All randomized patients accounted for in the analysis with one loss to follow-up.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Dougados 2013	High Risk	"Efficacy analyses were conducted in the intention-to-treat population (all randomly assigned and treated patients analysed in the arm they were randomly assigned to) with non-responder imputation for categorical variables (eg, DAS28 remission, ACR response), last observation carried forward until patient withdrawal for missing joint counts and no additional imputation of missing values."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Edwards 2004	Low Risk	Reasons documented for dropouts. For patients who withdrew before week 24, a last observation carried forward method of imputation was applied	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Emery 2006a	High Risk	Overall, 339 patients were randomized to three groups in the trial. 48 pf the 119 patients dropped out of the MTX+PL group. 31 of the 105 patients randomized to the Abatacept 2mg/kg + MTX group and 25 of the 115 patients randomized to the Abatacept 10mg/kg + MTX group dropped out. The reasons for discontinuations are not provided.	High Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Emery 2006b	High Risk	86% of abatacept group and 66% of placebo completed 6 months of treatment. At 12 months (results reported in Kremer 2005), 78% of abatacept group and 60% of placebo completed the study. Missing data were imputed for analysis. "Patients who discontinued the study because of worsening disease were considered to have had no response; for those who discontinued the study for other reasons the values for the last efficacy observation were carried forward." All withdrawals accounted for. Efficacy outcomes reported for total number of randomized patients. Judged as high risk of bias due to > 20% drop-out rate at 12 months in the treatment group.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Emery 2010	Low Risk	Over 90% of patients completed 48 weeks of the study	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Fleischmann 2003	High Risk	The study did not have full intention-to-treat analysis with results analyzed for only those who received at least one dose of the study drug (15 patients not included because of this). There was 78.1% follow-up through the entire study.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imabalances
Fleischmann 2009	High Risk	"All efficacy analyses were performed on the modified intent to treat (mITT) population (all randomised patients who had taken > 1 dose of study medication). The actual number of subjects in the summaries varies slightly from the mITT numbers due to non-imputable missing data for each parameter."	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imabalances
Furst 2003	Low Risk	A total of 578 (90.9%) patients completed 24 weeks of treatment, with no differences between the 2 groups	Unclear	Not enough information provided to assess selective reporting	Low Risk	"Demographic and baseline disease characterisitics were balanced between the groups at baseline."
Gashi 2014	low risk	No missing data	Unclear	Insufficient information	Unclear	Insufficient information
Genovese 2004	Low Risk	There was 83.6% follow-up through the entire study.	Unclear	Not enough information provided to assess selective reporting	Low Risk	Characterisitics were balanced across treatment groups
Genovese 2008	High Risk	"Efficacy analyses were conducted using the intent-to-treat population (ITT), which included all randomized patients who received 1 infusion of study treatment. Safety analyses included all randomized patients who received 1 infusion of study medication and had 1 postrandomization safety assessment. Patients who did not have the required data for a specific time point, who withdrew from the study, or who received rescue therapy were classified as nonresponders." Adequate intention to treat analysis requires that all patients who were randomized, not only those who received one dose of the study drug, to be included in the final analysis of the study.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Genovese 2011	High Risk	Efficacy data were assessed for both the per-protocol and the intent-to-treat (ITT) populations. The ITT population includes all randomized patients who received at least 1 dose of medication; the per-protocol population excludes patients with protocol violations." ITT is adequate when all patients randomized are included in the final analysis, not only those who receved one dosage of the drug.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Goekoop-Rulterman 2007	High Risk	The authors state that they performed an intention to treat analysis but there is no description of their analysis methods.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Heimans 2013	Low Risk	ITT analysis performed	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalance between two treatment arms
Huizinga 2015	Low Risk	"Efficacy analyses were performed in the intent-to-treat (ITT) population (randomised patients who received ≥ 1 administration of study medication) with non-responder imputation of missing data used for binary response variables (eg, DAS28-ESR remission and ACR response)." (p.37). Comment: Withdrawals due to AEs in the tocilizumab group were 8.7% compared to 9.1% in the placebo group. Total withdrawals were 19.9% in the tocilizumab group compared to 27.2% in the placebo group. 553/556 randomised patients were included in the ITT population. 3 people were omitted since they did not receive Tocilizumab.	Unclear	Comment: The protocol is not available.	Low Risk	No major baseline imbalances
Jobanputra 2012	High Risk	At 12 months, the authors provided data for the 39 and 34 patients still remaining in each study group respectively as well as a modified ITT population results. The modified method is described as follows: "Data for the modified intention to treat population with baseline values carried forward for those who discontinued therapy within 1 year."	Unclear	Not enough information provided to assess selective reporting	High Risk	Imbalances in concomitant treatments
Kaine 2011	Low Risk	De minimis dropouts in both groups	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kameda 2010	High Risk	"The primary analysis was conducted on an intention-to-treat population… Four patients were withdrawn from the study before the start of treatment, leaving 147 patients for the safety analysis (Fig. 1). Efficacy analysis was performed for 142 patients for whom sufficient clinical data were available. Sixteen patients (twelve in the E group and four in another) were withdrawn from the study by week 24 for the reasons given in Fig. 1." Insufficient data is presented as a reason why efficacy analysis was not performed on 2 patients in the ETN group and 3 patients in the ETN + MTX group. Moreover 12 more patients were lost in the ETN group than in ETN + MTX group.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kay 2008	High Risk	"For the primary analysis, a last observation carried forward procedure was used for patients who did not return for an evaluation or for whom we had insufficient data to determine their ACR20 response"	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Keystone 2004a	High Risk	"A total of 467 patients (75.4%) completed 52 weeks of treatment (Figure 1). Discontinuations occurred in 92 patients (22.0%) in the adalimumab groups and 60 patients (30.0%) in the placebo group."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Keystone 2004b	High Risk	"All patients who received at least 1 dose of study drug were included in the analyses of efficacy and safety…"	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Keystone 2008	Low Risk	"Efficacy analyses were conducted on an intent-to-treat (ITT) population, which consisted of all patients who were randomized into the study. Primary analyses were performed using nonresponder imputation. Patients who received rescue medication or who withdrew for any reason, including safety, were considered nonresponders from that time point onward."	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Keystone 2009	High Risk	No mention of whether ITT analysis was performed or how missing data was handled.	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kim 2007	Low Risk	" All 128 patients received at least one injection of the study drug and were thus included in the ITT analysis set."	Unclear	Not enough information provided to assess selective reporting	Low Risk	"Demographic characteristics were comparable between the two treatment groups."
Kim 2012	High Risk	"Efficacy analyses were performed on data from the modified intention-to-treat (mITT) population (defined as all randomized subjects who received at least one dose of study drug and had at least one post-baseline assessment) using the last observation carried forward approach, except for baseline values... All randomized subjects who received at least one dose of study medication were evaluated for safety (safety population)."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kim 2013	Low Risk	More than 80% completion	Unclear	Not enough information provided to assess selective reporting	High Risk	"Disease duration was longer in the PL group than in the INF group."
Kivitz 2014	high risk	High withdrawal and the withdrawals were imbalanced across the treatment groups	Unclear	insuffucent information	Low risk	Appears free of other biases
Kremer 2003	High Risk	86% of abatacept group and 66% of placebo completed 6 months of treatment. At 12 months (results reported in Kremer 2005), 78% of abatacept group and 60% of placebo completed the study. Missing data were imputed for analysis. "Patients who discontinued the study because of worsening disease were considered to have had no response; for those who discontinued the study for other reasons the values for the last efficacy observation were carried forward." All withdrawals accounted for. Efficacy outcomes reported for total number of randomized patients. Judged as high risk of bias due to > 20% drop-out rate at 12 months in the treatment group.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kremer 2005	High Risk	"All statistical analyses were carried out on the intent-totreat (ITT) population, defined as all patients who received at least 1 treatment infusion."	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kremer 2006	High Risk	89% of treatment group and 74% of placebo group completed the 1 year study. All withdrawals accounted for except for four patients that were randomized but not treated (Figure 1). "We performed all efficacy and safety analyses on a modified intention-to-treat population, defined as all randomly assigned patients who received at least 1 dose of study medication." However, "Nine abatacept-treated patients and 5 placebo recipients from 1 site were excluded from all efficacy analyses before unblinding due to nonadherence but were included in all safety analyses." Judged as high risk of bias due to exclusion of patients from efficacy analysis.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Kremer 2010	High Risk	"Efficacy data from all randomized patients were analyzed by assigned treatment group using an intent-to-treat approach…" However, patients assigned to lower dosages or inactive control were allowed to escape at 16 weeks to higher dosage or active treatment.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kremer 2011	Low Risk	"The ITT population comprised 1,190 patients; 86% of these patients completed 52 weeks, including 29% who received rescue therapy."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Kremer 2012	Low Risk	Discontinuations in each treatment group ranged from 12% to 22%, and the majority were not attributed to the study drug	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Lan 2004	High Risk	"The efficacy analysis was performed in the intent-totreat (ITT) population, which was defined as all patients with baseline data and at least 1 post-treatment evaluation. The last observation carried forward (LOCF) method was used to substitute for missing data."	High Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Lipsky 2000	High Risk	High rates of discontinuation in methotrexate group (50%) and infliximab group (21%).	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Lisbona 2008	Low Risk	No withdrawals or drop-outs.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Lisbona 2010	Low Risk	No withdrawals or drop-outs.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Machado 2014	Low risk	Proportion of missing outcomes compared with observed event risk not enough to induce clinically relevant bias in intervention effect estimate	Unclear	Insufficient information	Low risk	No major baseline imbalances
MacIsaac 2014	Low risk	No missing data	Unclear	insufficient information	Low risk	Appears free of other biases
Maini 1998	High Risk	More than 20% discontinued treatment	Unclear	Not enough information provided to assess selective reporting	High Risk	"There were some between-group differences in demographic characterisitics, baseline medication, or baseline disease activity, but these did not reach statistical significance, with the exception of the baseline HAQ Disability Index"
Maini 2006	High Risk	"Three hundred fifty-nine patients were randomly allocated to the 7 treatment groups. Patient flow through the trial and randomization to each treatment arm are shown in Figure 1. Approximately equal numbers of patients were randomly allocated to each arm of the study. During the trial, 60 patients withdrew (34 patients withdrew due to adverse events and/or possible drug-related toxicity). Of the 359 patients randomized to receive study medication, all were included in the safety population, and 354 patients were included in the full-analysis set (5 patients were excluded from the full-analysis set because of a protocol violation)."	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Moreland 2012	Low Risk	Statistical analysis was performed using the intention-to-treat approach	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Nishimoto 2009	High Risk	"The primary end point was the ACR20 response at week 24 with the last observation carried forward (LOCF) method, using an intent-to-treat (ITT) analysis…. All patients receiving at least one dose of tocilizumab or tocilizumab placebo, and at least 4 weeks of MTX or MTX placebo administration were included in the clinical efficacy analysis."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
O'Dell 2013	Low Risk	More than 80% completion	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Ohta 2014	Low risk	No missing data	Unclear	Insuffucent information	Low risk	Appears free of other biases

Pavelka 2009	High Risk	"Intent-to-treat analyses with the last observation carried forward (LOCF) were used to compare the two treatment arms."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Pope 2014	High risk	High withdrawal rate and per protocol analysis was used	Unclear	Insufficient information	Low risk	No major baseline imbalances
Rau 2004	Low Risk	No missing outcome data	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Rubbert-Roth 2010	High Risk	"Missing data were imputed using the non-responder method for ACR and EULAR (all patients who withdrew were classed as non-responders); last observation carried forward was used for all other endpoints. Further exploratory analyses were conducted to compare the 2 500mg group with the 2 1000mg group."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Schiff 2008	Low Risk	94.2% in abatacept group and 97.3% in placebo group completed treatment at 6 months. Missing data were imputed using LOCF or non-responders for ACR response. All withdrawals accounted for. "All patients who received at least one dose of study medication were assessed for efficacy and safety". Judged as low risk of bias given the > 80% completion rate in the treatment group.	Unclear	Not enough information provided to assess selective reporting	Low Risk	"Baseline demographics and clinical characteristics were similar between groups."
Schiff 2013	Low Risk	"The ITT analysis population includes all randomised patients who received at least one dose of medication."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Smolen 2008	High Risk	"The primary efficacy analysis was done on the intention-to-treat population (ITT)—ie, all patients randomised who received at least one infusion of study drug. The safety population included all randomised patients who received at least one infusion of study medication and who had at least one assessment of safety after randomisation. Patients who withdrew before week 24, patients who received rescue therapy, and patients whose week 24 categorical endpoints could not be determined due to insuffi cient data were deemed to be non-responders in the analysis. Last observation carried forward (LOCF) was used for tender and swollen joint counts; no imputation was used for missing HAQ, CRP, ESR, and global VAS data."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Smolen 2009	Unclear	No mention of whether ITT analysis was performed or how missing data was handled.	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Smolen 2013	High Risk	"In the open-label period, the modified intention-to-treat and safety populations included all patients who received at least one dose of etanercept. The radiographic population included all those who received at least one dose of study drug and had assessable baseline and post baseline radiographs. In the double-blind period, the modified intention-to-treat population was made up of patients who had received at least one dose of study drug and had one or more DAS28 evaluations. The safety population included all patients given at least one dose of the assigned treatment."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Smolen 2015	Low Risk	Missing values for components of efficacy endpoints are imputed, using median group values as determined by the patient's methotrexate stratification (yes/no) at baseline and last-observation-carried-forward methodology at all other time points." (additional file 1). "Efficacy data from one North American site that enrolled 16 patients were excluded because of protocol violations identified during standard audit processes. Patient baseline and safety data from these patients were not excluded." (p.3). Comment: 81/431 (18.8%) of patients discontinued due to adverse events.	Unclear	Comment: The protocol is not available.	Low Risk	No major baseline imbalances
Strand 2006	Low Risk	"At all time points, the proportion of patients continuing protocol participation was greatest in the rituximab+ methotrexate group—98%, 95% and 70% at weeks 24, 48 and 72, respectively."	Unclear	Not enough information provided to assess selective reporting	Low Risk	"Baseline demographics and disease characteristics were comparable across treatment groups."
Strand 2012	Unclear	No data to determine	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Tanaka 2011	High Risk	"All randomized patients received efficacy. To handle missing data in calculating response rates, 3 types of imputation were employed: nonresponder imputation, last observation carried forward (LOCF), and data as is (no imputation). LOCF was used for the primary analysis at week 12 for ACR20, and additional analyses were performed as a measure of robustness of the results."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Tanaka 2012	High Risk	"Efficacy and pharmacology parameters were primarily assessed according to a modified intent-to-treat approach in which patients who did not meet the study eligibility criteria, did not receive study treatment and/or had no effi cacy- or pharmacology- related data following randomisation were excluded from the full analysis patient population."	Unclear	Not enough information provided to assess selective reporting	High Risk	"Baseline demographic and disease characteristics were generally consistent across the three treatment groups, with the exception of shorter mean disease duration (8.1 years) and lower mean baseline CRP level (1.5 mg/dl) in Group 3 compared with Group 1 (8.7 years and 2.2 mg/dl, respectively) and Group 2 (8.8 years and 1.9 mg/dl, respectively)."
Taylor 2004	Low Risk	"No patient discontinued participation in the study during the first 18 weeks of therapy. One patient discontinued after week 18 because of lack of efficacy."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Taylor 2006	High Risk	"By week 110, a total of 5 patients had withdrawn from the study. Three of these patients consented to attend scheduled imaging procedures, and imaging data for these patients were included in the analyses. Thus, at week 110, data were available for 11 patients in each treatment group."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van der Heidje 2006	Low Risk	More than 80% completion	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van der Heijde 2007	High Risk	Less than 80% completion	High Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van der Heijde 2013	Low Risk	More than 80% completion	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van der Kooij 2009	Unclear	The authors state that they performed an intention to treat analysis but there is no description of their analysis methods.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van Riel 2006	Low Risk	"One patient in the ETN group did not receive drug and was withdrawn from the study, leaving 314 patients in the intent-to-treat analysis."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Van Riel 2008	High Risk	Patient reported outcomes were analyzed from Riel 2006 trial. The number of patients included in analysis for HAQ did not match those that were originally randomized to the treatment arms.	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van Vollenhoven 2012a	Low Risk	"The intention-to-treat population for this trial consisted of those patients who had received a treatment assignment. A few patients who did receive this assignment did not receive a dose of the allocated drugs (five patients in group A and eight patients in group B). Such patients are commonly excluded from the intention-to-treat population in a modified intention-to-treat analysis. However, because of the non-blinded nature of this trial, we chose to keep this small group of patients in the analysis, classifying them as non-responders."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Van Vollenhoven 2012b	High Risk	"The full analysis set for efficacy and safety included all patients who underwent randomization and who received at least one dose of study drug."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Weinblatt 1999	Low Risk	Clear descriptions of reasons for withdrawal of dropouts	Unclear	Not enough information provided to assess selective reporting	Low Risk	The treatment groups were generally well matched.
Weinblatt 2003	High Risk	"Among the 271 patients that entered the study, 161 completed the 24 weeks."	Unclear	Not enough information provided to assess selective reporting	Low Risk	There were no statistically significant differences in the demographic characteristics and baseline disease activity between the treatment groups
Weinblatt 2006	Unclear	More than 80% completed the study	Unclear	Not enough information provided to assess selective reporting	High Risk	CRP higher in subgroup receiving nonbiologic background therapy
Weinblatt 2012	Low Risk	More than 80% completion	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Weinblatt 2013a	High Risk	"Patients with prohibited medication usage, who discontinued because of lack of efficacy before week 14 and who lacked all week 14 ACR20 component data for any reason were considered ACR20 non-responders at week 14 through week 24. In these intent-to-treat analyses, patients randomised to placebo who EE (n = 68/197), and who received golimumab 2 mg/kg infusions at weeks 16 and 20, had week 16 data carried forward for response calculations at weeks 20 and 24. A last-observation-carried-forward procedure was employed to impute missing ACR component data (eg, swollen or tender joint count, or global assessments of disease) at week 14 if the patient."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Weinblatt 2013b	Low Risk	"At 1 year, 86.2% (274 of 318) of the SC abatacept–treated patients and 82% (269 of 328) of the SC adalimumab–treated patients completed the study."	Unclear	Not enough information provided to assess selective reporting	Low Risk	The demographic and clinical characteristics of the patients at baseline were balanced across the SC abatacept and SC adalimumab groups and were typical for RA studies (Table 1).
Weisman 2003	High Risk	"Efficacy was summarized on an intent-to-treat basis, including all patients who took 21 dose of study medication and having 21 available measurement."	High Risk	Not enough information provided to assess selective reporting	High Risk	"There was some variability in characteristics between treatment groups"
Weisman 2007	High Risk	"Forty-eight sites in the United States enrolled and randomized 564 patients; 535 patients received at least one dose of blinded study medication and were analysed."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances

Westhovens 2006	High Risk	"Data on all patients who received at least 1 infusion of study medication were included in the safety analysis, including the primary end point analysis, and were included in the treatment group that most closely corresponded to the infliximab dosage actually received. All efficacy analyses were performed using the intention-to-treat convention...The last observation was carried forward for missing data points."	Low Risk	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Yamamoto 2014a	Low risk	Missing data imputed using last observation carried forward (LOCF) non-responder imputation	Unclear	Insufficient information	Low risk	No major baseline imbalances
Yazici 2012	High Risk	"The intent-to-treat (ITT) population consisted of all randomly assigned patients who received at least one administration of study medication. The safety population included all patients who received at least one administration of study medication and who had at least one safety assessment after receiving study medication. Patients who received rescue therapy and patients who did not have data required to assess effi cacy outcomes at week 24 were classified as non-responders. Last observation carried forward methodology was used for missing joint count data."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances
Zhang 2006	High Risk	"Ten patients withdrew from the trial because of the development of adverse events: six in the infliximab group and four in the placebo group. In the infliximab group, two were due to skin rash, two because of leucopenia, one for pharyngitis, one because of the development of adult onset Still's disease. The patient who was diagnosed with adult onset Still's disease developed high fever after the third infusion of infliximab. The diagnosis was confirmed after further investigations but the researcher did not consider this was related to the treatment of infliximab. In the placebo group, two had exacerbation of joint symptoms, one for pneumonia and one for abnormality in serum alanine transaminase. Most adverse events were transient and returned to normal before the end of the trial. Others became normal during follow-up."	Unclear	Not enough information provided to assess selective reporting	Low Risk	No major baseline imbalances