Supplementary material

Study	Year	Population	Total n⁰ of patients	Female (%)*	Disease duration (years)*	Age (years)*	Intervention			Comparator		Concomitant medication
							Type	Daily dose	PBO	Active comparator		
Greenwald 2010 [34] NCT00902486	2010	inadequate response to DMARDs	124	n.a	7 - 9	54 - 58	bari	4,7,10 mg - 1x	+	-		csDMARD
I4V-MC-JADA [32,65,66] NCT01185353	2015	inadequate response to MTX	301	71-87%	5.3 - 6.6	49 - 53	bari	1 mg, 2 mg, 4 mg, 8 mg - 1x	+	-		MTX
RA-BALANCE [31,67] NCT02265705	2020	inadequate response to MTX	290	73.1-87.6%	9.1 - 10.7	48.9 - 49.5	bari	4 mg - 1x	+	-		MTX
RA-BEACON [29,68] NCT01721044	2016	inadequate response to TNFi	527	79-84%	14	55 - 56	bari	2 mg, 4 mg - 1x	+	-		csDMARD
RA-BEAM [28,69] NCT01710358	2017	inadequate response to MTX	1305	76-78%	10	53 - 54	bari	4 mg - 1x	+	adalimumab		csDMARD
RA-BEGIN [59,70,71] NCT01711359	2017	csDMARD naive	584	70-76%	1.3 - 1.9	49 - 51	bari	4 mg - 1x	+	MTX		csDMARD
RA-BUILD [30,64,72] NCT01721057	2017	inadequate response to csDMARDs	684	80-83%	7 - 8	51 - 52	bari	2 mg, 4 mg - 1x	+	-		csDMARDs
Tanaka 2016 [33] NCT01469013	2016	inadequate response to MTX	145	71-92%	5.06 - 6.32	51.1 - 57.5	bari	1 mg, 2 mg, 4 mg, 8 mg - 1x	+	-		MTX
Fleischmann 2015 [53,73] NCT01052194	2015	inadequate response to MTX	204	78-85%	6.3 - 10	54.9 - 56.8	decerno	25 mg, 50 mg, 100 mg, or 150 mg - 2x	+	-		
Genovese 2016 [51,74] NCT2011-004419-22	2016	inadequate response to MTX	358	71.8-87.3%	6.5-8.1	50.1 - 53.5	decerno	100 mg, 150 mg, 200 mg - 1x, or 100 mg - 2x	+	-		MTX
Genovese 2016 II [52] NCT01754935	2016	inadequate response to DMARDs	43	50-83.3%	6.8 - 11	50.5 - 56.7	decerno	100 mg, 200 mg, 300 mg - 1x	+	-		DMARDs
DARWIN 1 [46] NCT01888874	2016	inadequate response to MTX	594	76.5-86%	7 - 10	52 - 55	filgo	50, 100 or 200 mg 1x, 50, 100 or 200 mg - 2x	+	-		MTX
DARWIN 2 [45] NCT01894516	2017	inadequate response to MTX	283	75.7-87%	9 - 10	52 - 53	filgo	50, 100 or 200 mg - 1x,	+	-		
FINCH 1 [43] NCT02889796	2019	inadequate response to MTX	1755	n.a	n.a	n.a	filgo	100 mg, 200 mg - 1x	+	adalimumab		MTX
FINCH 2 [44] NCT02873936	2019	inadequate response or intolerance to bDMARDs	448	77.8-81.8%	9.8 - 10.3 (median)	55 - 56	filgo	100 mg, 200 mg - 1x	+	-		csDMARD
FINCH 3 [57] NCT02886728	2019	MTX naive	1249	n.a	n.a	n.a	filgo	100 mg, 200 mg, 200 mg (mono) - 1x	+	MTX		
Vanhoutte 2017 I [62] NCT01384422	2017	n.a	36	91.7%	5.6 - 9.7	47 - 53	filgo	200 mg - 1x or 100 mg - 2x	+	-		MTX
Vanhoutte 2017 II [62] NCT01668641	2017	n.a	91	72.7-82.3%	4.4 - 10	44 - 55	filgo	30 mg, 75 mg, 150 mg, or 300 mg - 1x	+	-		MTX
Luchi [63] NCT01626573	2013	n.a	60	73%	n.a	53.2	ita	100 mg, 200 mg - 2x, 300 mg, 60 mg - 1x	+	-		csDMARDs
Genovese 2017 [47,75] NCT01565655	2017	inadequate response to csDMARDs	289	78-87.9%	9.8-10	52.6 - 54.9	pefi	25 mg, 50 mg, 100 mg, 150 mg - 1x	+	-		
Kivitz 2016 [48,76] NCT01554696	2017	inadequate response to MTX	379	81-87.5%	7.2 - 8.1	52.3 - 54.5	pefi	25 mg, 50 mg, 100 mg, or 150 mg - 1x	+	-		MTX

Study	Year	Inclusion criteria	N	%		Age	Drug	Dose		Comparator	Background
RAJ3 [50] NCT02308163	2019	inadequate response to csDMARDs	507	69-76.5%	6.98 - 10.39	54.1 - 56.3	pefi	100 mg, 150 mg - 1x	+	etanercept	MTX
RAJ4 [49] NCT02305849	2019	inadequate response to MTX	519	67.8-71.8%	4.3 - 4.41	55.3 - 58.5	pefi	100 mg, 150 mg - 1x	+	-	MTX
Takeuchi 2016 [60] NCT01649999	2016	no restrictions	281	76.4-83.6%	6.92 - 8.03	51.6 - 54.2	pefi	25, 50, 100, 150 mg - 1x	+	-	
Robinson 2020 [54] NCT02969044	2020	inadequate response to MTX	70	78.6-85.7%	7.1-8.4	54.2 - 55.4	ritle	200 mg - 1x	+	-	MTX
Boyle 2015 [23] NCT00976599	2015	Inadequate response to MTX	29	85.7-93.3%	5.5 - 12.2	53.1 - 53.5	tofa	10 mg - 2x	+	-	MTX
Conaghan 2016 [55] NCT01164579	2016	MTX naive	109	78.4-86.1%	0.6 - 0.8	47.8 - 50.8	tofa	10 mg - 2x	+	MTX	
Fleischmann 2012 [21] NCT00550446	2012	inadequate response to DMARDs	384	85.2-88.1%	8.1 - 10.8	52 - 55	tofa	1mg, 3 mg, 5 mg, 10 mg, 15 mg - 2x	+	adalimumab	
Kremer 2009 [24,77] NCT00147498	2009	inadequate response to DMARDs	264	84.1-87%	8.7 - 10.2	47.9 - 51.8	tofa	5 mg, 15 mg, 30 mg - 2x	+	-	
Kremer 2012 [25] NCT00413660	2012	inadequate response to DMARDs	507	74.3-88%	7.5 - 11.8	51 - 56	tofa	1mg, 3 mg, 5 mg, 10 mg, 15 mg - 2x, 20 mg - 1x	+	-	MTX
Menshikova 2018 [61] n.a	2018	n. a	30	70%	3.5	48.2	tofa	5 mg - 2x	-	etanercept	n. a
Nakamura 2018 [22] NCT02157012	2018	inadequate response to DMARDs	50	59-85%	3.4 - 3.5	67.2 - 68.3	tofa	5 mg - 2x	-	tocilizumab, abatacept	MTX
ORAL Scan [15,78] NCT00847613	2013	Inadequate response to MTX	797	80.2 - 91.1%	8.8 - 9.5	52.0 - 53.7	tofa	5 mg, 10 mg - 2x	+	-	MTX
ORAL Solo [21,79] NCT00814307	2012	inadequate response to DMARDs	610	85.2-88.2%	7.7 - 8.6	49.7 - 52.4	tofa	5 mg, 10 mg - 2x	+	-	
ORAL Standard [16,80,81] NCT00853385	2012	Inadequate response to MTX	717	75.0 - 85.3%	6.9 - 9.0	51.9 - 55.5	tofa	5 mg, 10 mg - 2x	+	adalimumab	MTX
ORAL Start [58,82,83] NCT01039688	2014	MTX naive	956	76.7-82.4%	2.9-3.4	48.8 - 50.3	tofa	5 mg, 10 mg - 2x	-	MTX	
ORAL Step [18,84] NCT00960440	2013	inadequate response to TNFi	399	80.3-86,6%	11.3–13.0	54.4 - 55.4	tofa	5 mg, 10 mg - 2x	+	-	MTX
ORAL Strategy [20,85,86] NCT02187055	2017	Inadequate response to MTX	1146	83%	5.4 - 6.1	49.7 - 50.7	tofa	5 mg - 2x	-	adalimumab	MTX
ORAL Sync [17,87] NCT00856544	2013	inadequate response to DMARDs	792	75.0 - 83.8%	8.1 - 10.2	50.8 - 53.3	tofa	5 mg, 10 mg - 2x	+	-	nonbiologic DMARDs
Tanaka 2011 [26] NCT00603512	2011	inadequate response to MTX	136	75-96.2%	5.7 - 8.7	50 - 53.3	tofa	1mg, 3 mg, 5 mg, 10 mg - 2x	+	-	MTX
Tanaka 2015 [27] NCT00687193	2015	inadequate response to at least one synthetic or bDMARD	317	79.2-88.7%	6.8 - 11	52.6 - 54.7	tofa	1mg, 3 mg, 5 mg, 10 mg, 15 mg - 2x	+	-	
BALANCE I [39] NCT01960855	2016	inadequate response to TNFi	276	78-86%	10.9 - 12.3	56 - 59	upa	3 mg, 6 mg, 12 mg, or 18 mg - 2x	+	-	MTX
BALANCE II [38] NCT02066389	2016	inadequate response to MTX	299	68-86%	3.9 - 9.3	53 - 56	upa	3 mg, 6 mg, 12 mg, or 18 mg - 2x, or 24 mg - 1x	+	-	MTX
SELECT-BEYOND [40,88,89] NCT02706847	2018	inadequate response to bDMARDs	499	84-85%	12.4 - 14.5	56.3 - 57.6	upa	15 mg, 30 mg - 1x	+	-	csDMARD
SELECT-CHOICE [41] NCT03086343	2020	inadequate response to bDMARDs	612	81.9-82.2%	11.8 - 12.4	55.3 - 55.8	upa	15 mg - 1x	-	abatacept	csDMARD
SELECT-COMPARE [36,86] NCT02629159	2019	inadequate response to MTX	1629	79-80%	8	54	upa	15 mg - 1x	+	adalimumab	MTX

Study	Year	Population	N	%	Duration*	Age*	Drug	Dose			
SELECT-EARLY [56,90,91] NCT02706873	2020	MTX naive	945	76-76.4%	2.6-2.9	51.9 - 54.9	upa	15 mg, 30 mg - 1x	-	MTX	
SELECT-MONOTHERAPY [35,92] NCT02706951	2019	inadequate response to MTX	648	79-83%	5.8 - 7.5	53.1 - 55.3	upa	15 mg, 30 mg - 1x	-	MTX	
SELECT-NEXT [37,93] NCT02675426	2018	inadequate response to csDMARDs	661	75-82%	7.2 - 7.3	55.3 - 56	upa	15 mg, 30 mg - 1x	+	-	csDMARD
SELECT-SUNRISE [42] NCT02720523	2020	inadequate response to csDMARDs	197	69.4-86	2.1 - 4 (median)	54.3 - 56	upa	7,5 mg, 15 mg, 30 mg - 1x	+	-	csDMARD

Supplementary Table S1. Characteristics of the included studies

All included studies are RCTs. *Values are means, ranging among study groups, unless stated otherwise. DMARD denotes disease modifying antirheumatic drugs; csDMARD, conventional synthetic DMARD; bDMARD, biological DMARD; TNFi, tumor necrosis factor inhibitor; MTX, methotrexate; tofa, tofacitinib; bari, baricitinib; upa, upadacitinib; filgo, filgotinib; pefi, peficitinib; decerno, decernotinib; ritle, ritlecitinib, ita, itacitinib and n.a, not available

Supplementary material

Summary of outcomes

ACR20 response   Number of patients who reached at least 20% improvement in measures involved in ACR Core Data Set (SJC, TJC, PtGA, PGA, pain, ESR or CRP)

ACR50 response   Number of patients who reached at least 50% improvement in measures involved in ACR Core Data Set (SJC, TJC, PtGA, PGA, pain, ESR or CRP)

ACR70 response   Number of patients who reached at least 70% improvement in measures involved in ACR Core Data Set (SJC, TJC, PtGA, PGA, pain, ESR or CRP)

CDAI difference   LSM change from baseline in the score on CDAI

CDAI remission   Number of patients reaching remission defined by CDAI ($\leq$ 2.8)

CRP   LSM change from baseline in CRP value (mg/l)

DAS28-CRP difference   LSM change from baseline in the score on DAS28-CRP

DAS28-CRP remission   Number of patients reaching remission defined by DAS28-CRP (< 2.6)

DAS28-ESR difference   LSM change from baseline in the score on DAS28-ESR

DAS28-ESR remission   Number of patients reaching remission defined by DAS28-ESR (< 2.6)

Deaths   Number of patients who died during study period

Discontinuation   Number of patients experiencing symptoms leading to discontinuation of the study drugs

EQ-5D (US)   LSM change from baseline in the score on EQ-5D assessed by the US scoring algorithm

EQ-5D (US)	LSM change from baseline in the score on EQ-5D assessed by the UK scoring algorithm
EQ-5D (VAS)	LSM change from baseline in the score on EQ-5D assessed by the VAS
ESR	LSM change from baseline in ESR value (mm/hour)
FACIT-F	LSM change from baseline in the score on FACIT-F
HAQ-DI difference	LSM change from baseline in the score on HAQ-DI
HAQ-DI improvement	Number of patients reaching at least the minimum clinically important difference (≥0.22 MCID) in HAQ-DI
MJS duration	LSM change from baseline in the duration of morning joint stiffness
MOS-Sleep	LSM change from baseline in the score on MOS-Sleep
Pain	LSM change from baseline in pain measured on VAS ranging from 0-100 mm
PGA	LSM change from baseline in PGA measured on VAS ranging from 0-100 mm
PtGA	LSM change from baseline in PtGA measured on VAS ranging from 0-100 mm
SDAI difference	LSM change from baseline in the score on SDAI
SDAI remission	Number of patients reaching remission defined by SDAI (≤ 3.3)
Serious side effects	Number of patients experiencing serious side effects during the study
SF-36 MCS	LSM change from baseline in the score on SF-36 assessing the mental component score

SF-36 PCS	LSM change from baseline in the score on SF-36 assessing the physical component score
Side effects	Number of patients experiencing side effects during the study
Swollen joint counts	LSM change from baseline in the number of swollen joint counts
Tender joint counts	LSM change from baseline in the number of tender joint counts
WPAI A	LSM change from baseline in the score on WPAI assessing absenteeism (missed worked time)
WPAI AI	LSM change from baseline in the score on WPAI assessing activity impairment
WPAI OWI	LSM change from baseline in the score on WPAI assessing overall work impairment (productivity loss)
WPAI P	LSM change from baseline in the score on WPAI assessing presenteeism (impairment while working)

Supplementary Table S2. Summary of the investigated outcomes

ACR denotes American College of Rheumatology; SJC, swollen joint counts; TJC, tender joint counts; PtGA, Patient's Global Assessment of Disease Activity; PGA; Physician's Global Assessment of Disease Activity; CDAI, Clinical Disease Activity Index; LSM, least square mean, CRP, C-reactive protein; DAS28-CRP, Disease Activity Score 28 using C-reactive protein; DAS28-ESR, Disease Activity Score 28 using erythrocyte sedimentation rate; EQ-5D UK, EuroQol 5 Dimensions Questionnaire - UK scoring algorithm; EQ-5D US, EuroQol 5 Dimensions Questionnaire - US scoring algorithm and EQ-5d (VAS), EQ-5D measured on Visual Analogue Scale; ESR, erythrocyte sedimentation rate; FACIT-F, Functional Assessment of Chronic Illness Therapy – Fatigue; HAQ-DI, Health Assessment Questionnaire - Disability Index; MJS, Morning Joint Stiffness; MOS-Sleep, Medical Outcomes Study Sleep Scale; SDAI, Simple Disease Activity Index; SF-36 MCS, 36-Item Short Form Survey – Mental Component

Score, SF-36 PCS, SF-36 Physical Component Score ; WPAI AI, Work Productivity and Activity Impairment Questionnaire – Activity Impairment; WPAI A, WPAI - Absenteeism; WPAI - P, WPAI Presenteeism; WPAI - OWI, WPAI Overall Work Impairment;

	JAK inhibitors vs Placebo	JAK inhibitors vs bDMARDS	JAK inhibitors vs MTX
ACR20 response	yes	yes	yes
ACR50 response	yes	yes	yes
ACR70 response	yes	yes	yes
CDAI difference	yes	no*	no*
CDAI remission	yes	yes	yes
CRP	yes	yes	no*
DAS28-CRP difference	yes	no*	no*
DAS28-CRP remission	yes	yes	yes
DAS28-ESR difference	yes	no*	no*
DAS28-ESR remission	yes	yes	yes
Deaths	yes	yes	no*
Discontinuation	yes	no*	no*
EQ-5D (US)	yes	no*	no*
EQ-5D (US)	yes	no*	no*
EQ-5D (VAS)	yes	no*	no*
ESR	yes	no*	no*
Erosion score	no*	no*	no*
FACIT-F	yes	yes	yes
HAQ-DI difference	yes	yes	yes

HAQ-DI Improvement	yes	no*	yes
Joint-space narrowing score	no*	no*	no*
MJS duration	yes	no*	no*
MJS severity	no*	no*	no*
MOS-Sleep	yes	no*	no*
Modified total Sharp score	no*	no*	no*
Pain	yes	no*	no*
Pain catastrophizing scale	no*	yes	no*
PGA	yes	yes	no*
PtGA	yes	yes	no*
SDAI difference	yes	no*	no*
SDAI remission	yes	yes	yes
Serious side effects	yes	yes	yes
SF-36 MCS	yes	yes	no*
SF-36 PCS	yes	yes	yes
Side effects	yes	no*	no*
Swollen joint counts	yes	yes	no*
Tender joint counts	yes	yes	no*
Worst Joint Pain	no*	no*	no*
Worst Tiredness	no*	no*	no*
WPAI A	yes	no*	no*
WPAI AI	yes	no*	no*
WPAI OWI	yes	yes	no*
WPAI P	yes	no*	no*

Supplementary Table S3. List of the investigated outcomes in each comparisons

*Outcomes could not be included in the quantitative analysis if (1) did not appear at least in 3 of the randomized controlled trials measured in the same way or (2) about which data only at a time-point with a not acceptable study design could be obtained.

Summary of risk of bias

Unique ID	Experimental	Comparator	Weight	Randomization process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall
Greenwald 2010	baricitinib	placebo	1	?	+	+	?	?	!
I4V-MC-JADA	baricitinib	placebo	1	+	+	+	+	+	+
RA-BALANCE	baricitinib	placebo	1	+	+	+	+	+	+
RA-BEACON	baricitinib	placebo	1	+	+	+	+	+	+
RA-BEAM	baricitinib	placebo and active agent	1	+	+	+	+	+	+
RA-BEGIN	baricitinib	placebo and active agent	1	+	+	+	+	+	+
RA-BUILD	baricitinib	placebo	1	+	+	+	+	+	+
Tanaka 2016	baricitinib	placebo	1	+	+	+	+	+	+
Fleischmann 2015	decernotinib	placebo	1	?	+	+	+	+	!
Genovese 2016	decernotinib	placebo	1	?	+	?	+	+	!
Genovese 2016 II	decernotinib	placebo	1	?	+	+	+	+	!
DARWIN 1	filgotinib	placebo	1	+	+	+	+	+	+
DARWIN 2	filgotinib	placebo	1	+	+	+	+	+	+
FINCH 1	filgotinib	placebo and active agent	1	+	+	+	+	+	+
FINCH 2	filgotinib	placebo	1	+	+	+	+	+	+
FINCH 3	filgotinib	placebo and active agent	1	+	+	+	+	+	+
Vanhouette 2017 I	filgotinib	placebo	1	+	+	+	+	+	+
Vanhouette 2017 II	filgotinib	placebo	1	+	+	+	+	+	+
Luchi 2013	itacitinib	placebo	1	?	+	+	+	?	!
Genovese 2017	peficitinib	placebo	1	?	+	+	+	+	!
Kivitz 2016	peficitinib	placebo	1	?	+	+	+	+	!
RAJ 3	peficitinib	placebo	1	+	+	+	+	+	+
RAJ 4	peficitinib	placebo	1	+	+	+	+	+	+
Takeuchi 2016	peficitinib	placebo	1	+	+	+	+	+	+
Robinson 2020	rilecitinib	placebo	1	+	+	+	+	+	+
Boyle 2015	tofacitinib	Placebo	1	?	+	+	+	+	!
Conaghan 2016	tofacitinib	placebo and active agent	1	+	?	?	+	+	!
Fleischmann 2012	tofacitinib	placebo and active agent	1	?	+	+	?	+	!
Kremer 2009	tofacitinib	Placebo	1	?	+	?	+	+	!
Kremer 2012	tofacitinib	placebo	1	?	+	?	+	+	!
Menshikova 2018	tofacitinib	active agent	1	?	?	−	?	?	−
Nakamura 2018	tofacitinib	active agents	1	?	?	+	?	?	!
ORAL Scan	tofacitinib	placebo	1	+	+	+	+	+	+
ORAL Solo	tofacitinib	placebo	1	+	+	+	+	+	+
ORAL Standard	tofacitinib	placebo and active agent	1	+	+	+	+	+	+
ORAL Start	tofacitinib	active agent	1	+	+	+	+	+	+
ORAL Step	tofacitinib	placebo	1	+	+	+	+	+	+
ORAL Strategy	tofacitinib	active agent	1	+	+	+	+	+	+
ORAL Sync	tofacitinib	placebo	1	+	+	+	+	+	+
Tanaka 2011	tofacitinib	placebo	1	+	+	+	+	+	+
Tanaka 2015	tofacitinib	placebo	1	+	+	+	+	+	+
BALANCE I	upadacitinib	placebo	1	+	+	+	+	+	+
BALANCE II	upadacitinib	placebo	1	+	+	+	+	+	+
SELECT-BEYOND	upadacitinib	placebo	1	+	+	+	+	+	+
SELECT-CHOICE	upadacitinib	active agent	1	+	+	+	+	+	+
SELECT-COMPARE	upadacitinib	placebo and active agent	1	+	+	+	+	+	+
SELECT-EARLY	upadacitinib	active agent	1	+	+	+	+	+	+
SELECT-MONOTHERAPY	upadacitinib	active agent	1	+	+	+	+	+	+
SELECT-NEXT	upadacitinib	placebo	1	+	+	+	+	+	+
SELECT-SUNRISE	upadacitinib	placebo	1	+	+	+	+	+	+

Supplementary Table S4. Summary of the overall risk of bias in the included studies
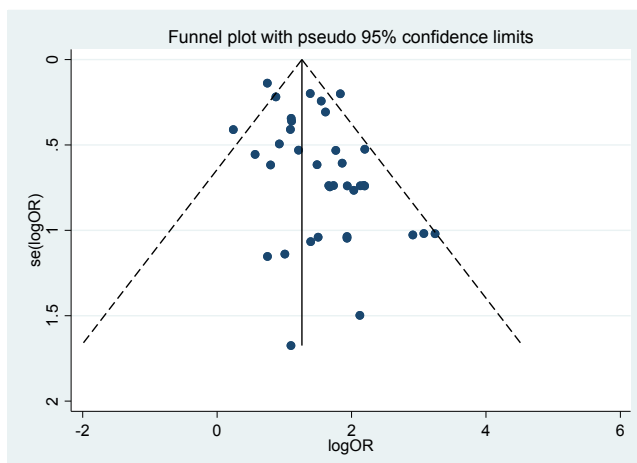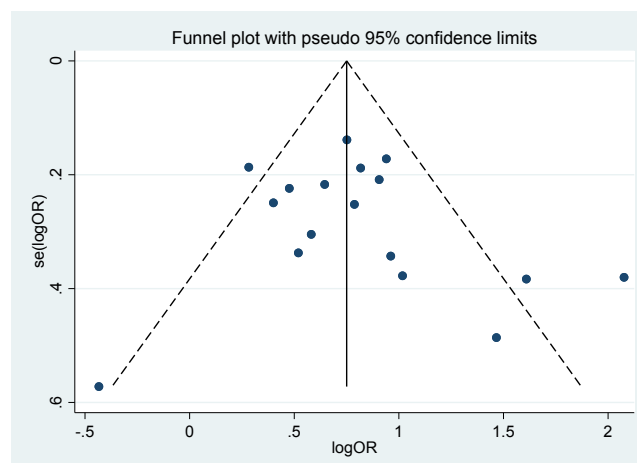
Summary of results
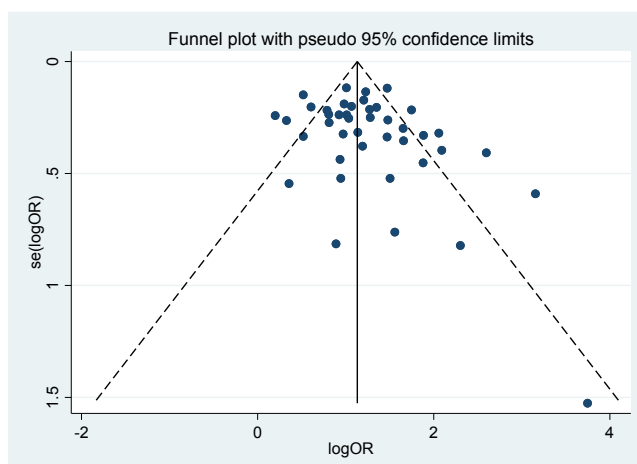
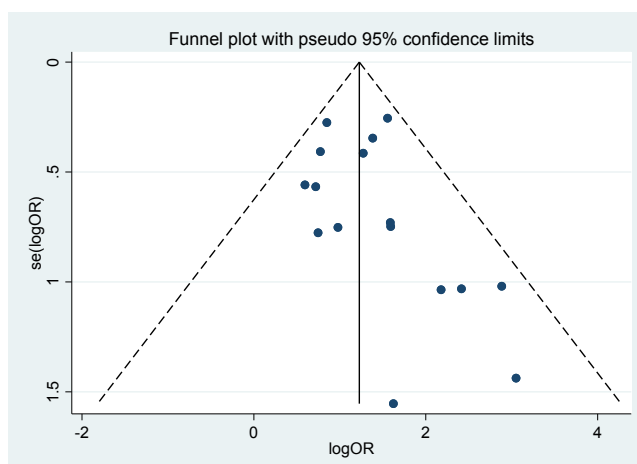S1. Funnel plot of ACR 20 response (JAK inhibitor vs placebo)
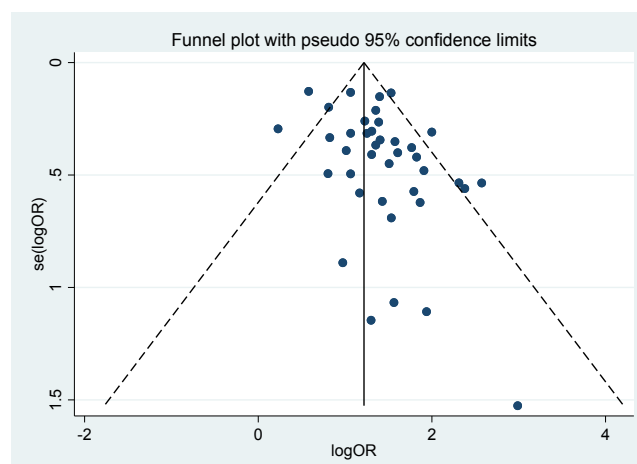
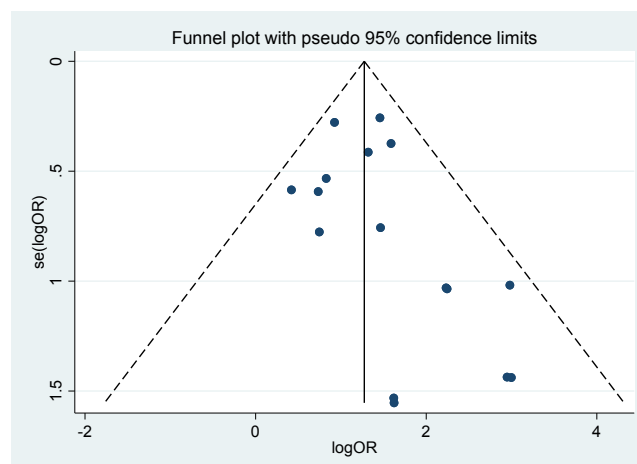S2. Funnel plot of ACR 50 response (JAK inhibitor vs placebo)

S3. Funnel plot of ACR 70 response (JAK inhibitor vs placebo)
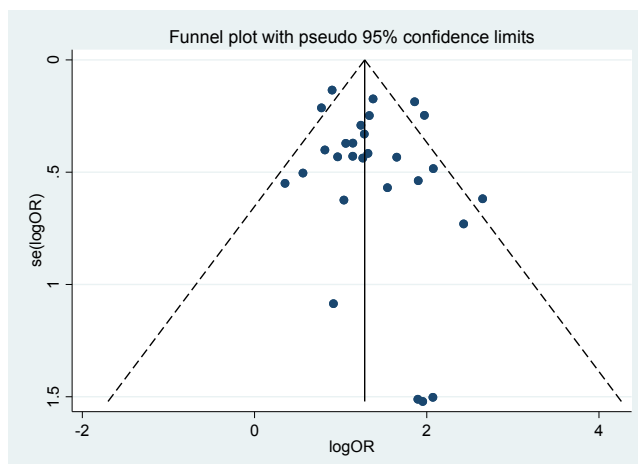

S4. Funnel plot of HAQ-DI improvement (JAK inhibitor vs placebo)

S5. Funnel plot of CDAI remission (JAK inhibitor vs placebo)

S6. Funnel plot of SDAI remission (JAK inhibitor vs placebo)

S7. Funnel plot of DAS28-CRP remission (JAK inhibitor vs placebo)
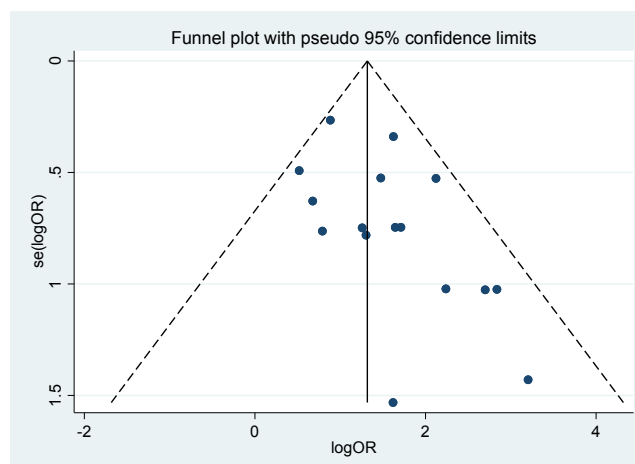
S8. Funnel plot of DAS28-ESR remission (JAK inhibitor vs placebo)
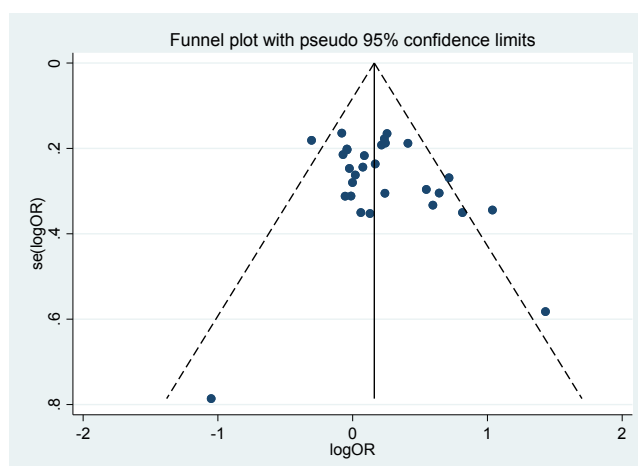
S9. Funnel plot of side effects remission (JAK inhibitor vs placebo)

S10. Funnel plot of serious side effects remission (JAK inhibitor vs placebo)
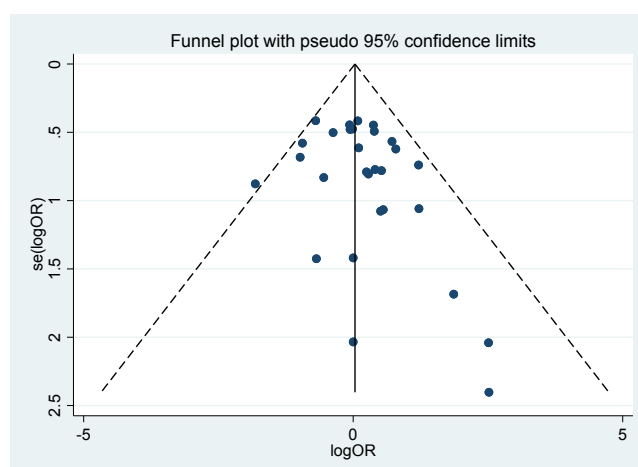
S11. Funnel plot of discontinuation (JAK inhibitor vs placebo)
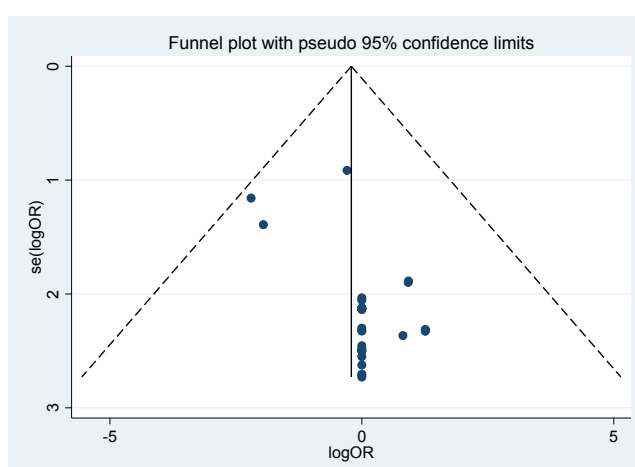
S12. Funnel plot of deaths (JAK inhibitor vs placebo)
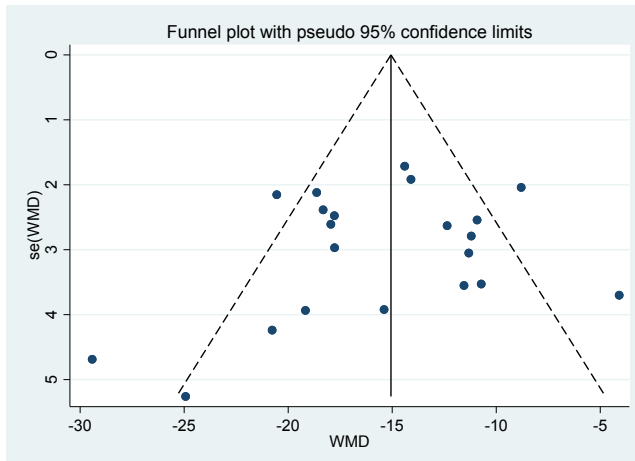
S13. Funnel plot of Pain VAS (JAK inhibitor vs placebo)
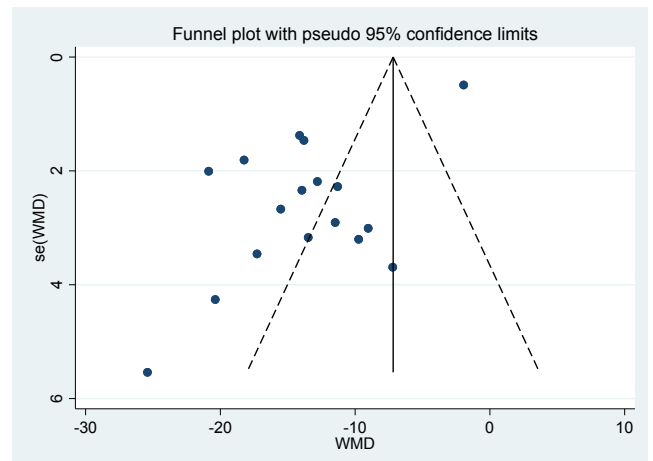
S14. Funnel plot of PGA VAS (JAK inhibitor vs placebo)
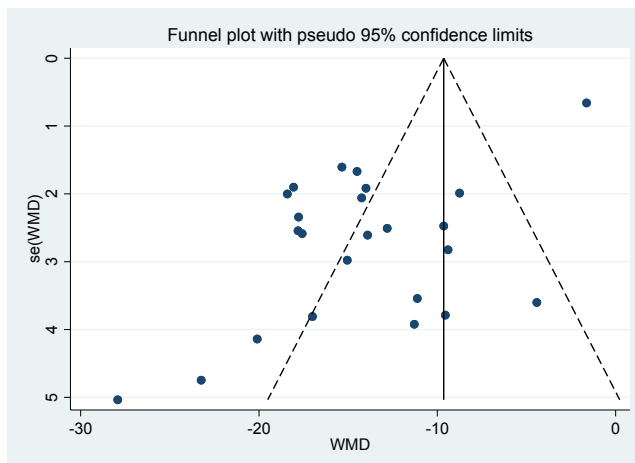
S15. Funnel plot of PtGA VAS (JAK inhibitor vs placebo)
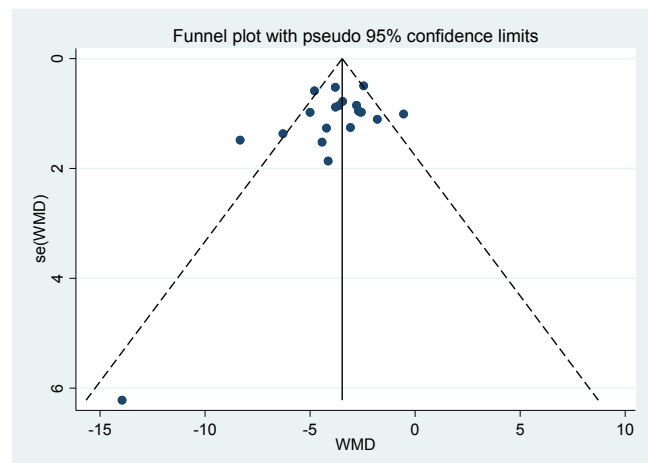
S16. Funnel plot of Swollen Joint Count (JAK inhibitor vs placebo)
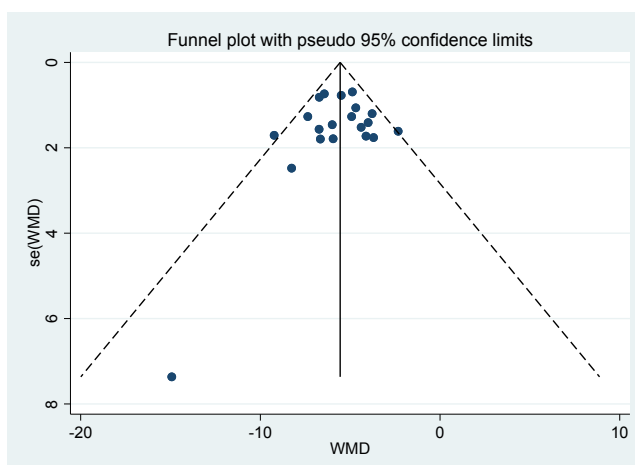
S17. Funnel plot of Tender Joint Count (JAK inhibitor vs placebo)
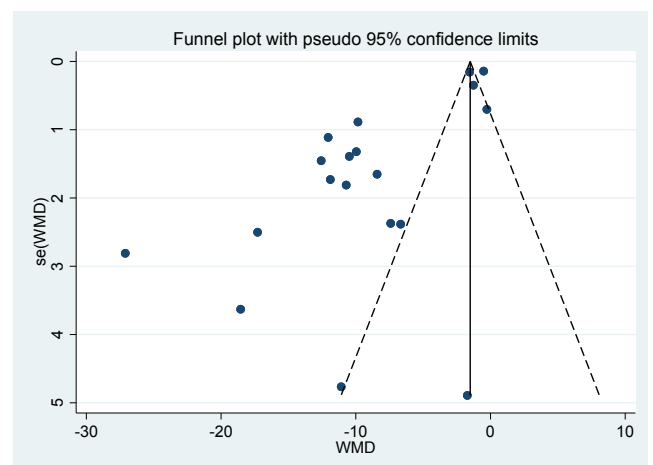
S18. Funnel plot of CRP (JAK inhibitor vs placebo)
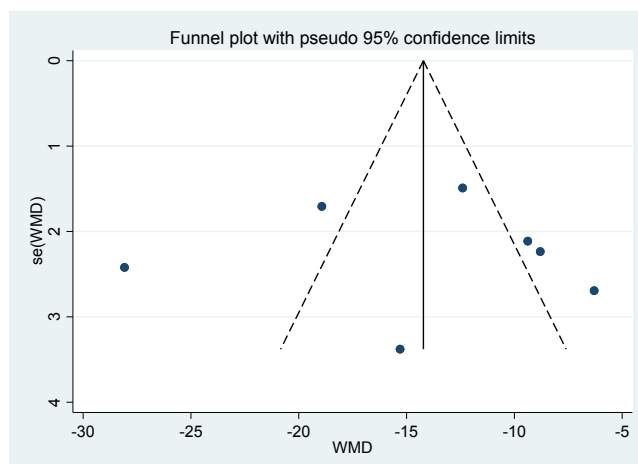
S19. Funnel plot of ESR (JAK inhibitor vs placebo)
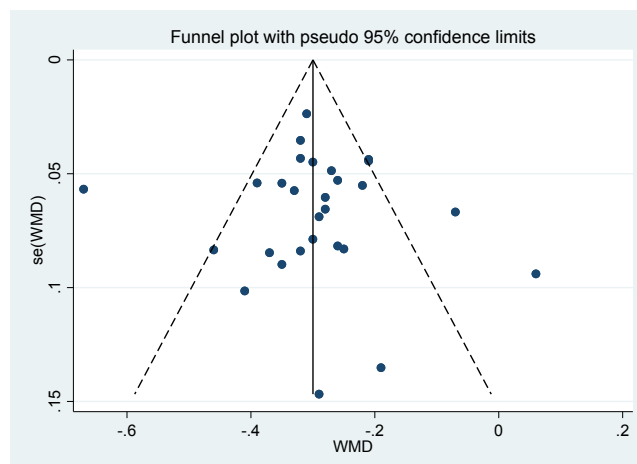
**S20. Funnel plot of HAQ-DI improvement (JAK inhibitor vs placebo)**
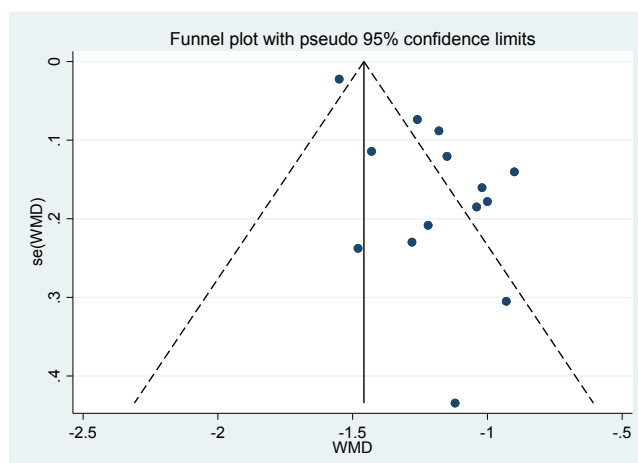
S21. Funnel plot of DAS28-CRP difference (JAK inhibitor vs placebo)
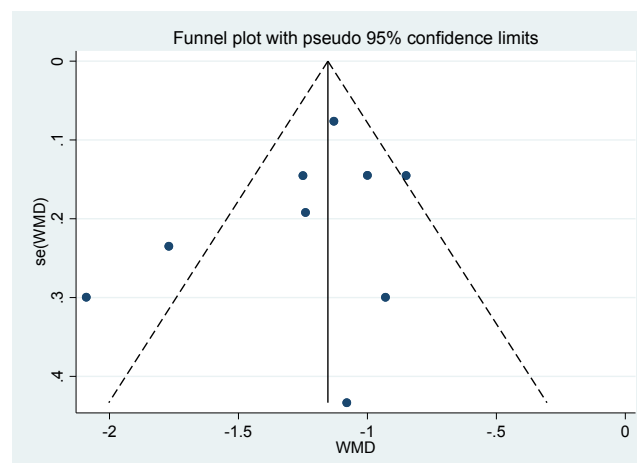
S22. Funnel plot of DAS28-ESR difference (JAK inhibitor vs placebo)
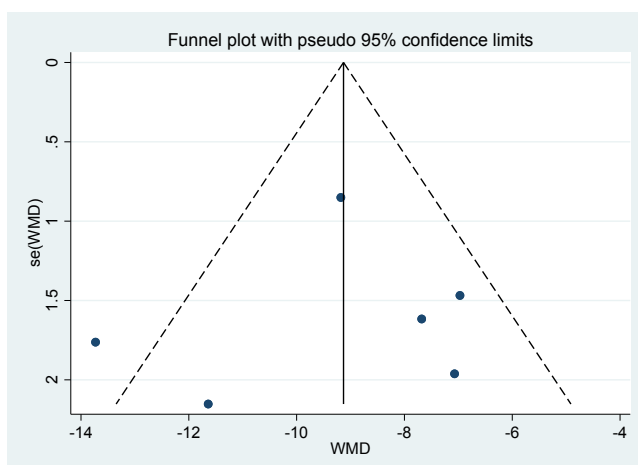
S23. Funnel plot of CDAI difference (JAK inhibitor vs placebo)
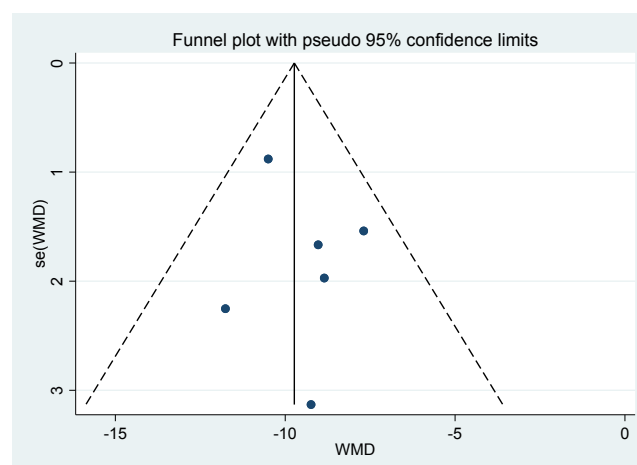
S24. Funnel plot of SDAI difference (JAK inhibitor vs placebo)
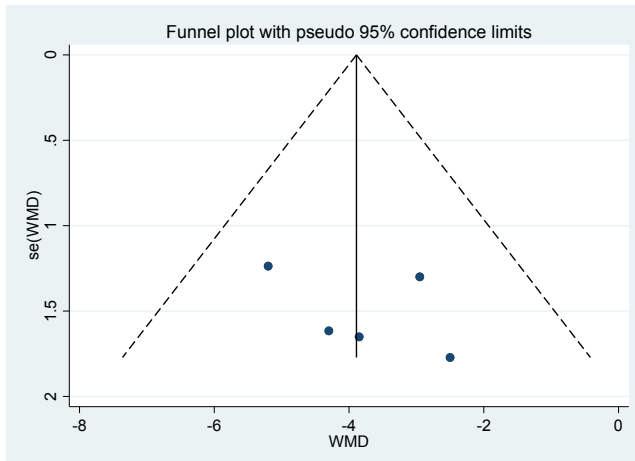
S25. Funnel plot of MOS-Sleep (JAK inhibitor vs placebo)
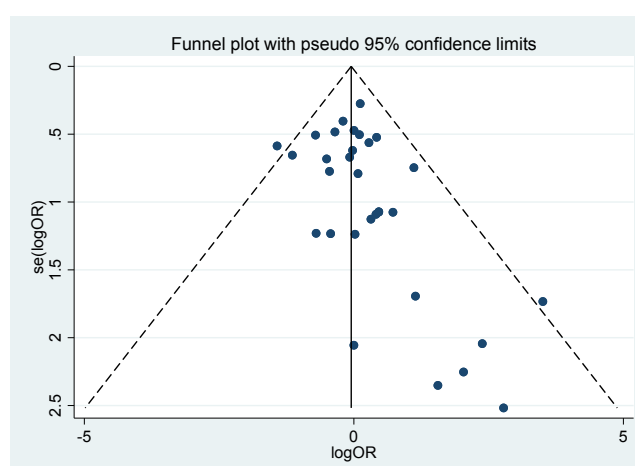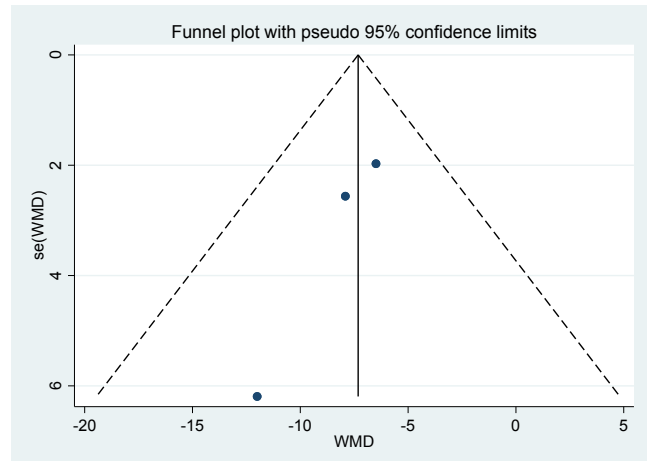

S26. Funnel plot of WPAI AI (JAK inhibitor vs placebo)
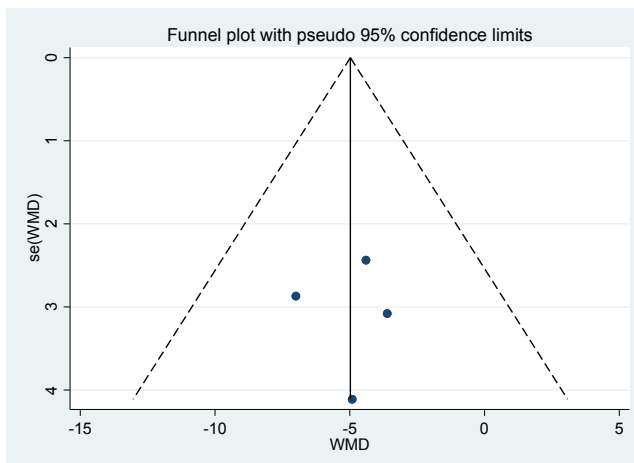
S27. Funnel plot of WPAI A (JAK inhibitor vs placebo)
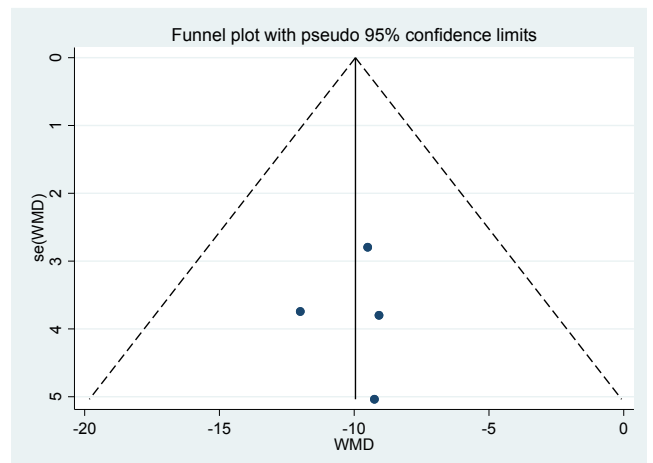
S28. Funnel plot of WPAI OWI (JAK inhibitor vs placebo)
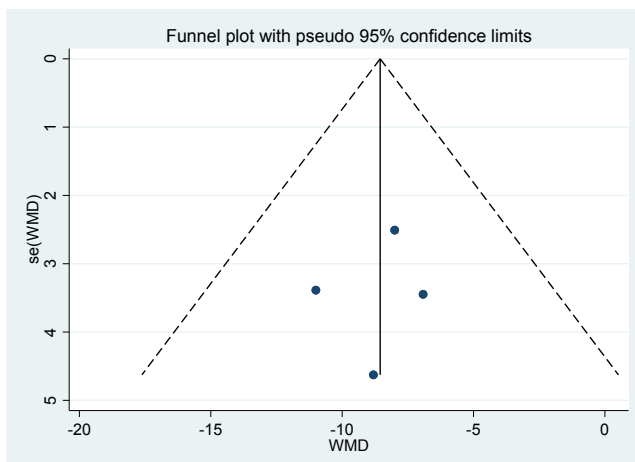
S29. Funnel plot of WPAI P (JAK inhibitor vs placebo)
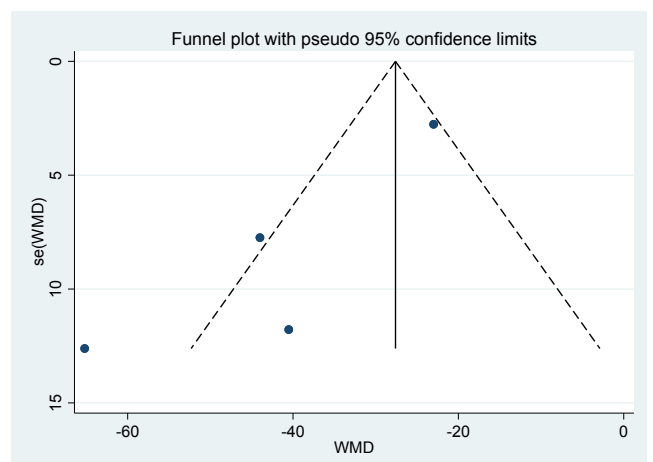
S30. Funnel plot of MJS duration (JAK inhibitor vs placebo)
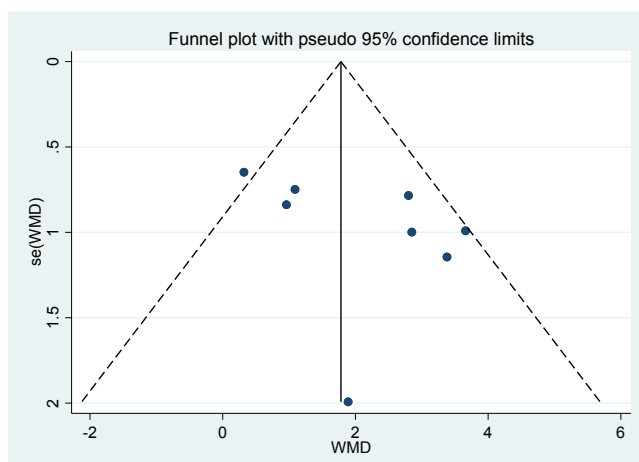
S31. **Funnel plot of SF-36 MCS (JAK inhibitor vs placebo)**
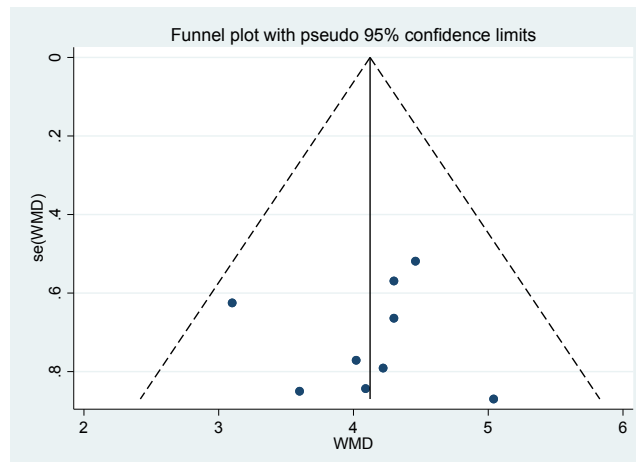
S32. **Funnel plot of SF-36 PCS (JAK inhibitor vs placebo)**
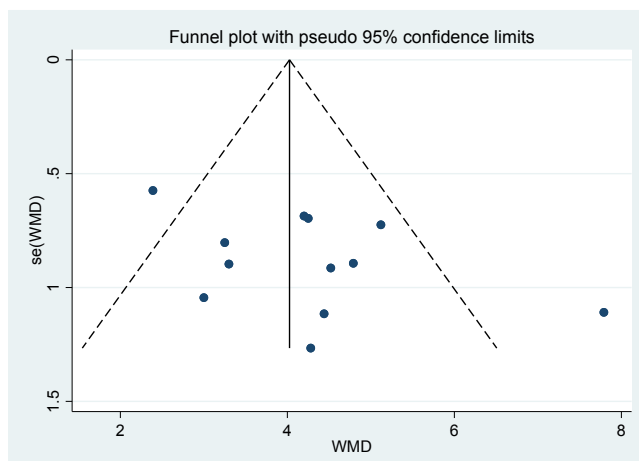
S33. **Funnel plot of FACIT-F (JAK inhibitor vs placebo)**
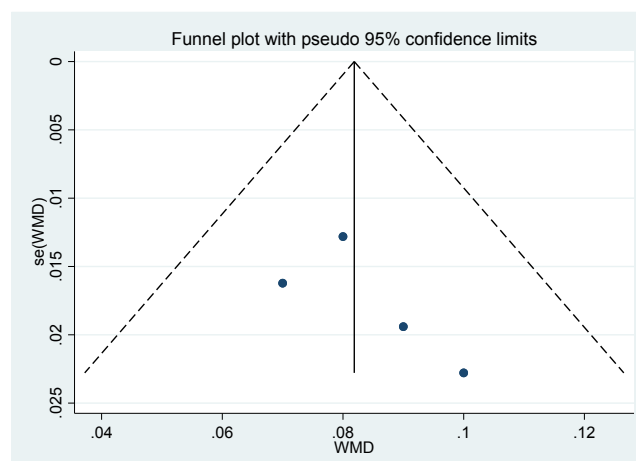
S34. **Funnel plot of EQ-5D UK (JAK inhibitor vs placebo)**
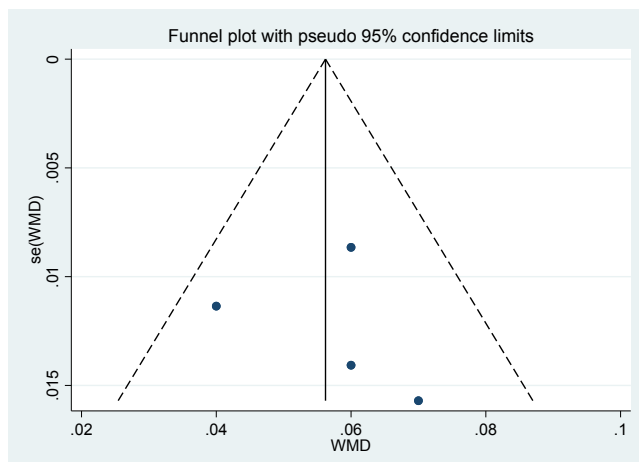
S35. **Funnel plot of EQ-5D US (JAK inhibitor vs placebo)**
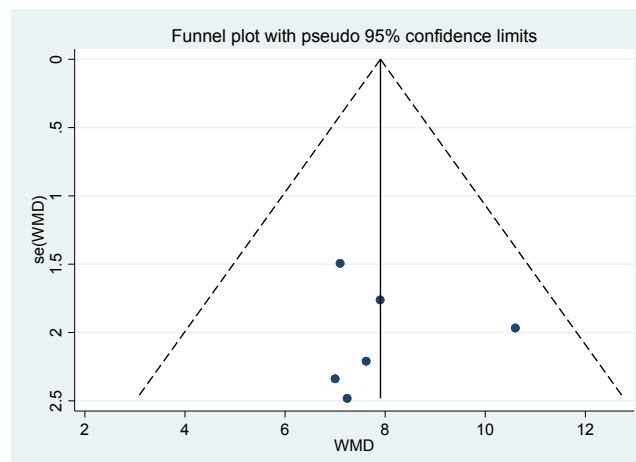
S36. **Funnel plot of EQ-5D VAS (JAK inhibitor vs placebo)**
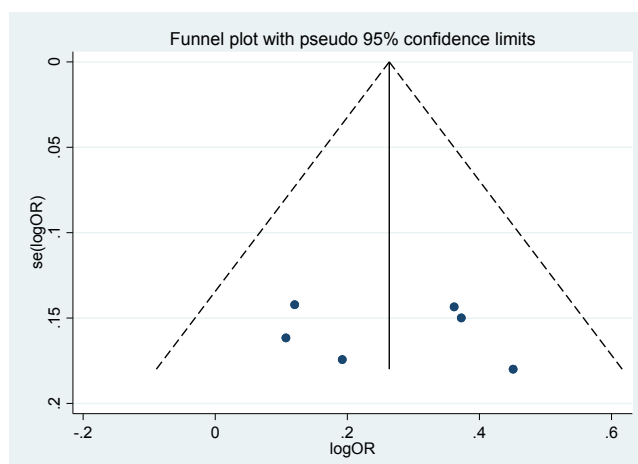
S37. Funnel plot of ACR20 response (JAK inhibitor vs bDMARDs)
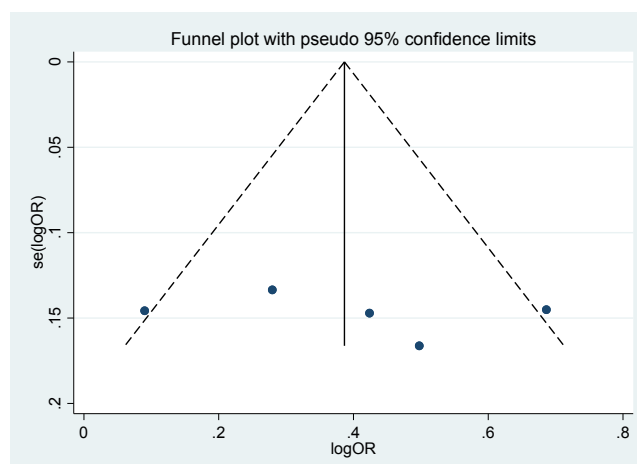
S38. Funnel plot of ACR50 response (JAK inhibitor vs bDMARDs)
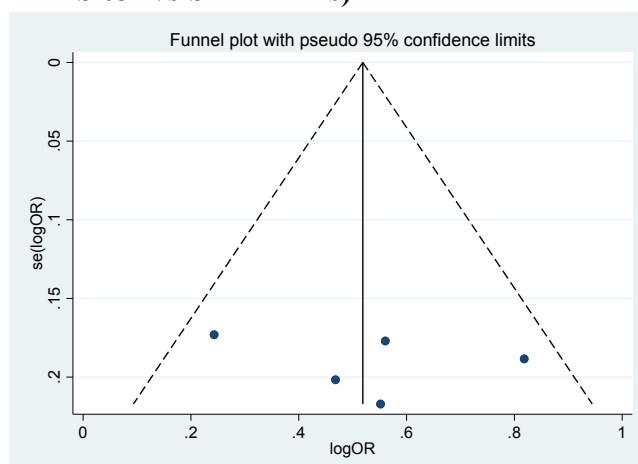
S39. Funnel plot of ACR70 response (JAK inhibitor vs bDMARDs)
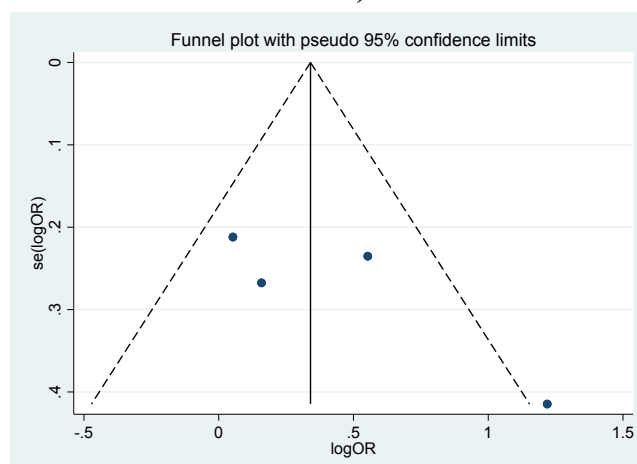
S40. Funnel plot of CDAI remission (JAK inhibitor vs bDMARDs)
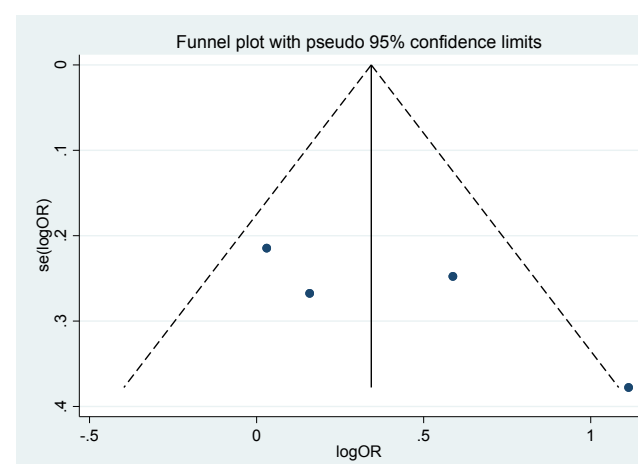
S41. Funnel plot of SDAI remission (JAK inhibitor vs bDMARDs)
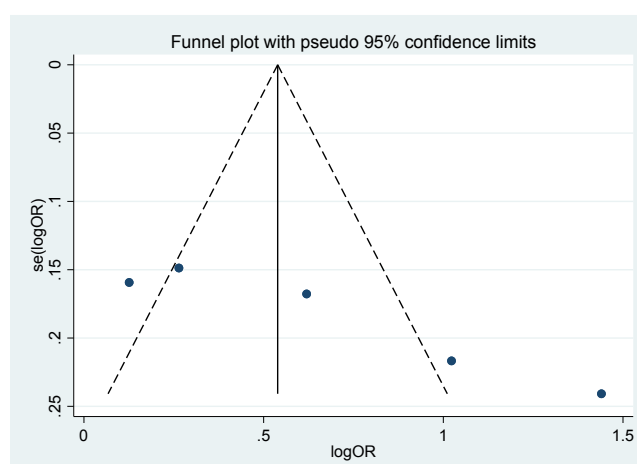
S42. Funnel plot of DAS28 - CRP remission (JAK inhibitor vs bDMARDs)
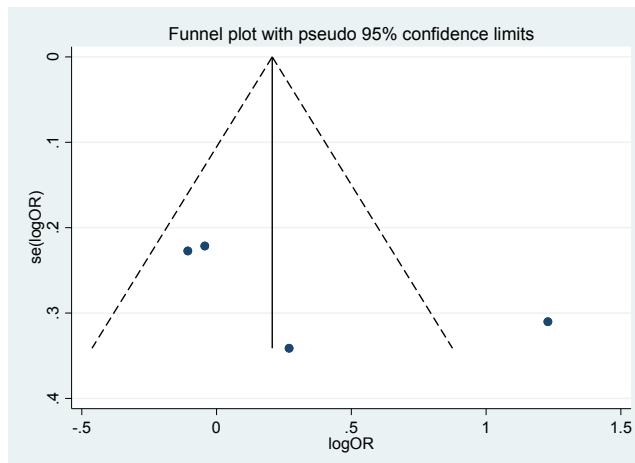
S43. Funnel plot of DAS28 - ESR remission (JAK inhibitor vs bDMARDs)
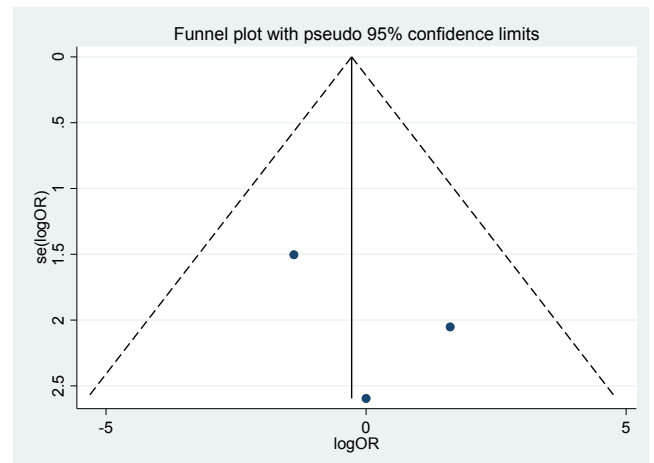
S44. Funnel plot of deaths (JAK inhibitor vs bDMARDs)
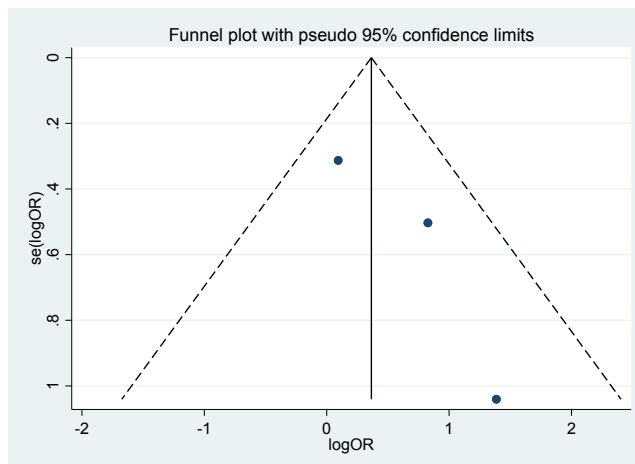
S45. Funnel plot of serious side effects (JAK inhibitor vs bDMARDs)
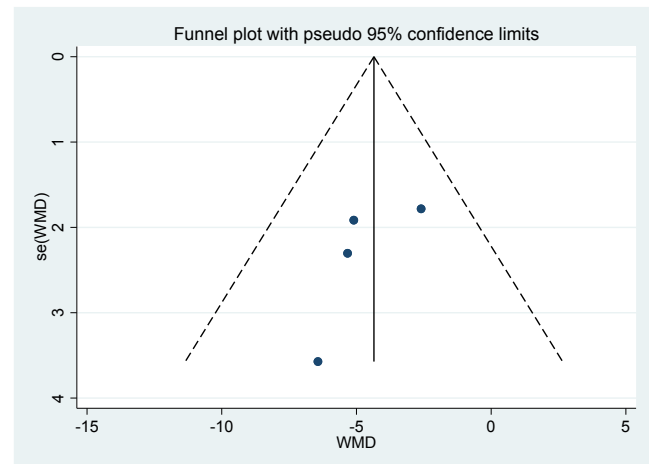
S46. Funnel plot of Pain VAS (JAK inhibitor vs bDMARDs)
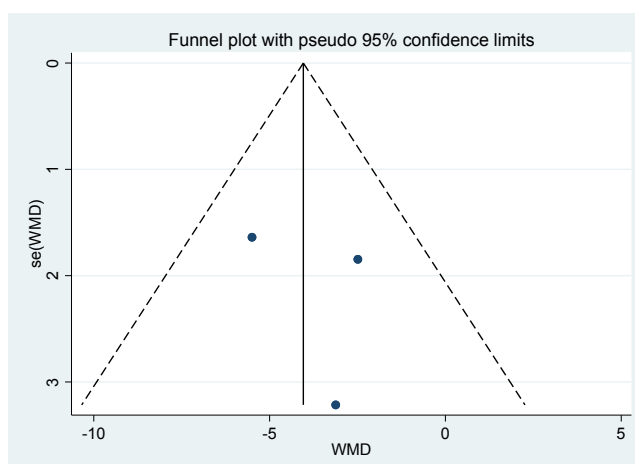
S47. Funnel plot of PGA VAS (JAK inhibitor vs bDMARDs)
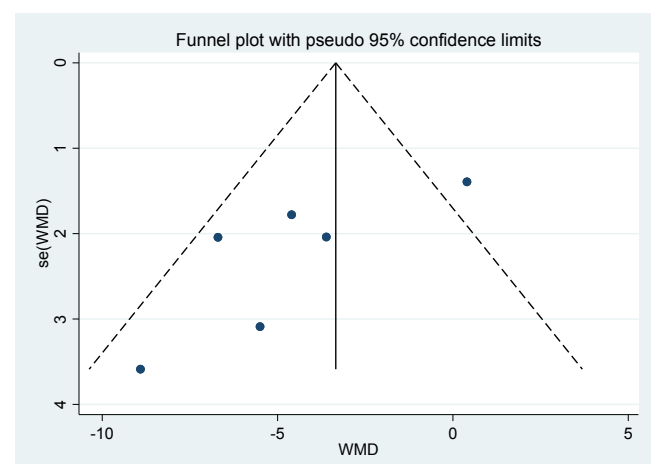
S48. Funnel plot of PtGA VAS (JAK inhibitor vs bDMARDs)
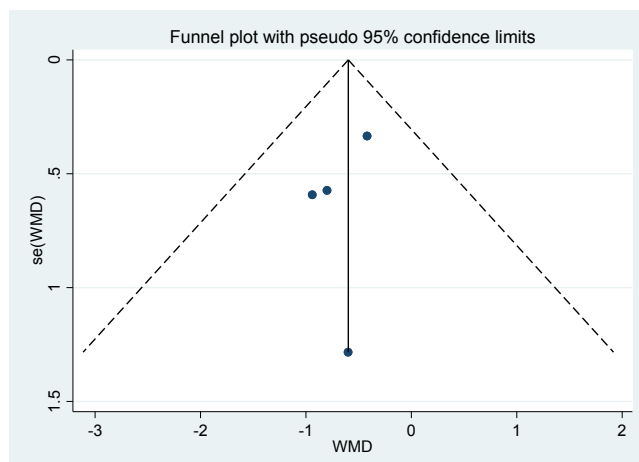
S49. Funnel plot of Swollen Joint Count (JAK inhibitor vs bDMARDs)
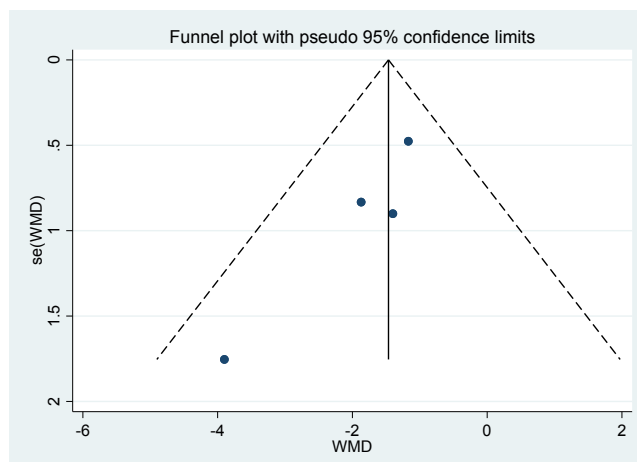
S50. Funnel plot of Tender Joint Count (JAK inhibitor vs bDMARDs)
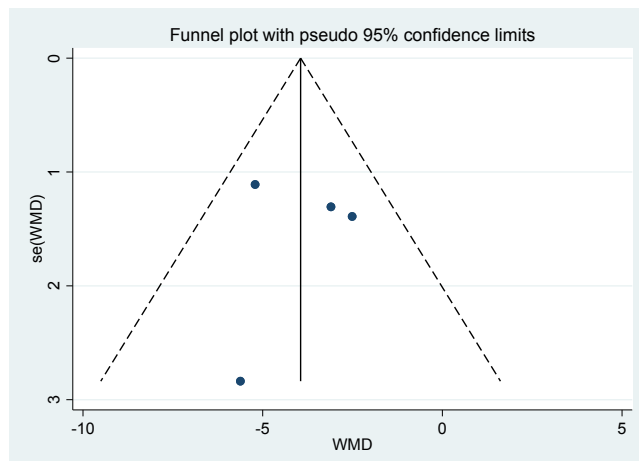
S51. Funnel plot of CRP (JAK inhibitor vs bDMARDs)
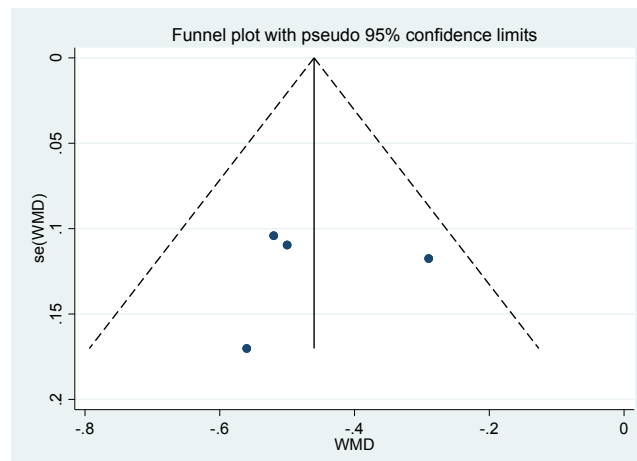
S52. Funnel plot of DAS28 - CRP difference (JAK inhibitor vs bDMARDs)
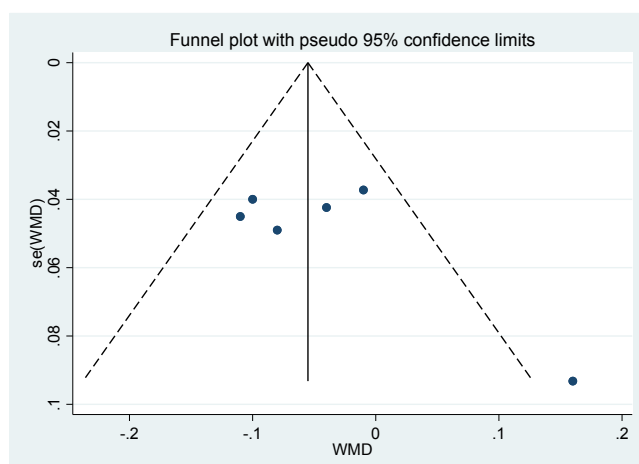
S53. Funnel plot of HAQ-DI difference (JAK inhibitor vs bDMARDs)
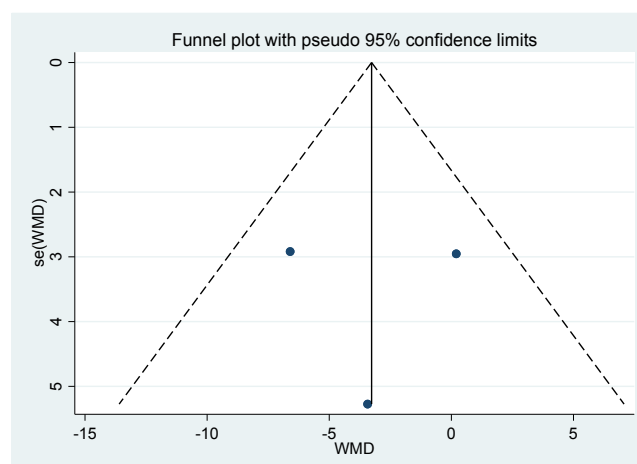
S54. Funnel plot of WPAI OWI (JAK inhibitor vs bDMARDs)
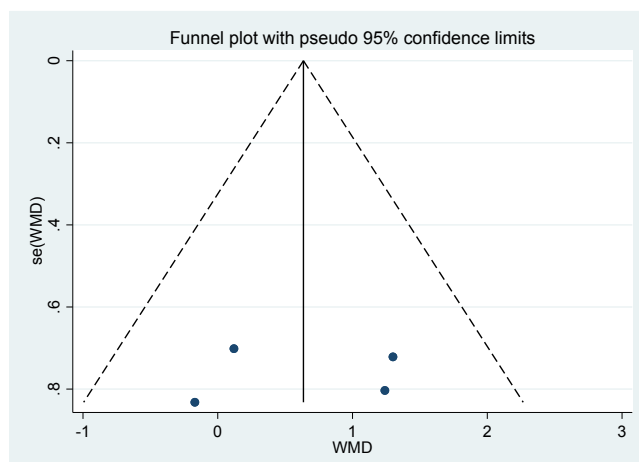
S55. Funnel plot of SF-36 MCS (JAK inhibitor vs bDMARDs)
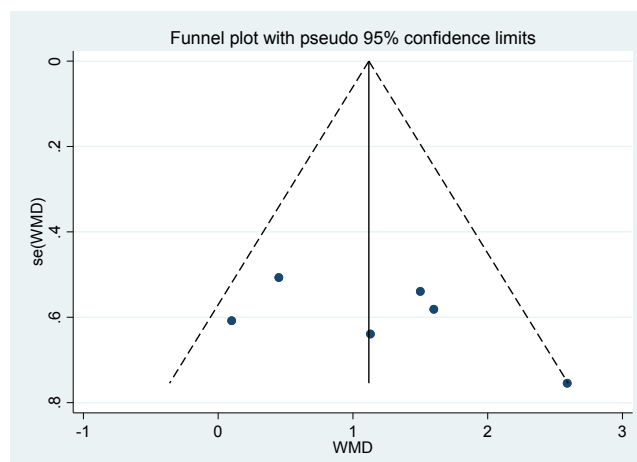
S56. Funnel plot of SF-36 PCS (JAK inhibitor vs bDMARDs)
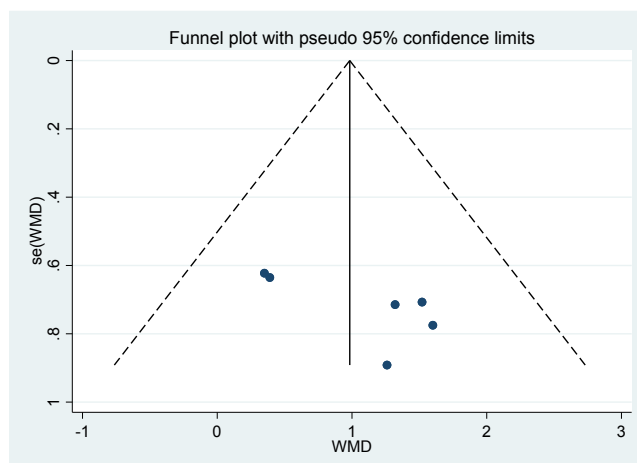
S57. Funnel plot of FACIT-F (JAK inhibitor vs bDMARDs)
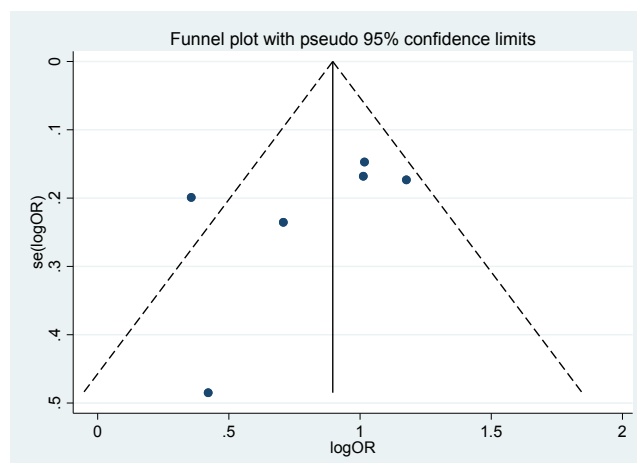
S58. Funnel plot of ACR20 response (JAK inhibitor vs MTX)
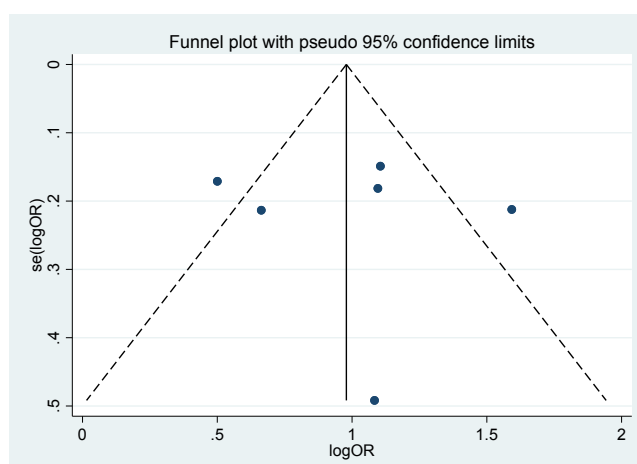
S59. Funnel plot of ACR50 response (JAK inhibitor vs MTX)
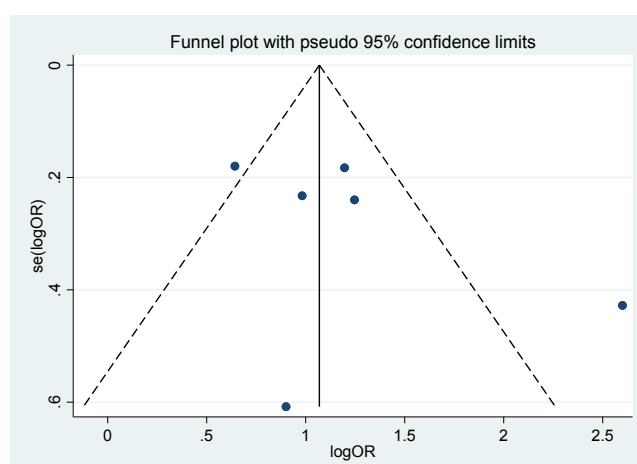
S60. Funnel plot of ACR70 response (JAK inhibitor vs MTX)
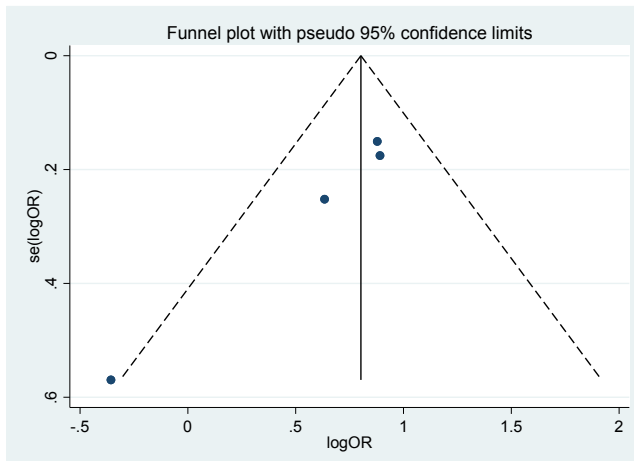
S61. Funnel plot of HAQ - DI improvement (JAK inhibitor vs MTX)
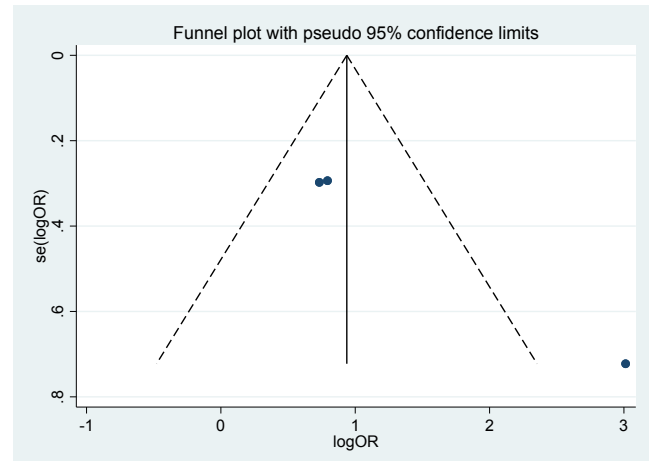
S62. Funnel plot of CDAI remission (JAK inhibitor vs MTX)
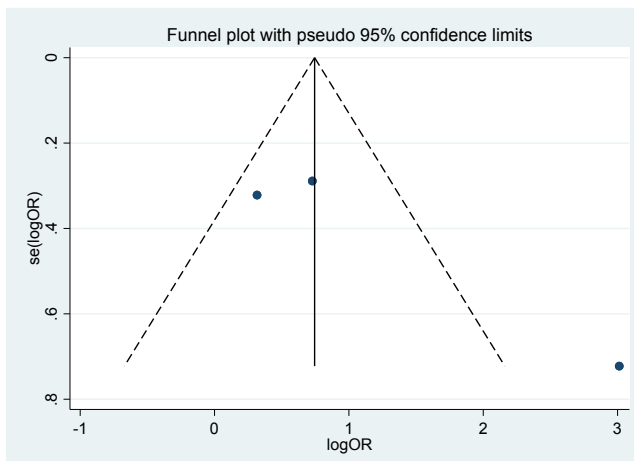
S63. Funnel plot of SDAI remission (JAK inhibitor vs MTX)
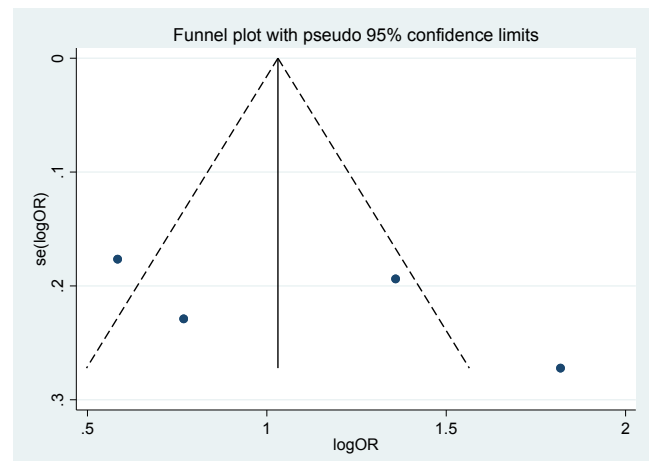
S64. Funnel plot of DAS28 – CRP remission (JAK inhibitor vs MTX)
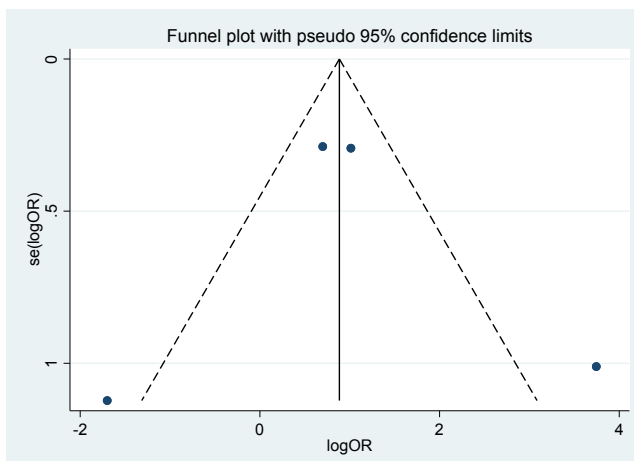
S65. Funnel plot of DAS28 – ESR remission (JAK inhibitor vs MTX)
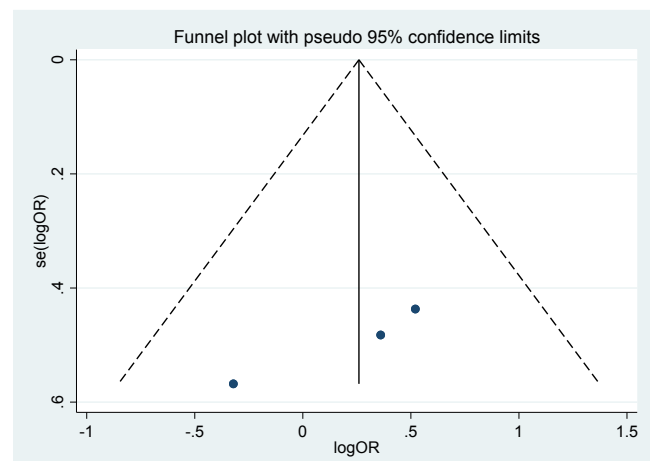
S66. Funnel plot of serious side effects (JAK inhibitor vs MTX)
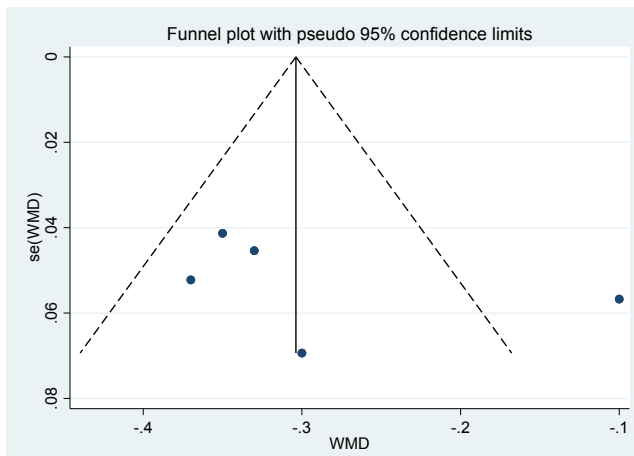
S67.   Funnel plot of HAQ-DI difference (JAK inhibitor vs MTX)
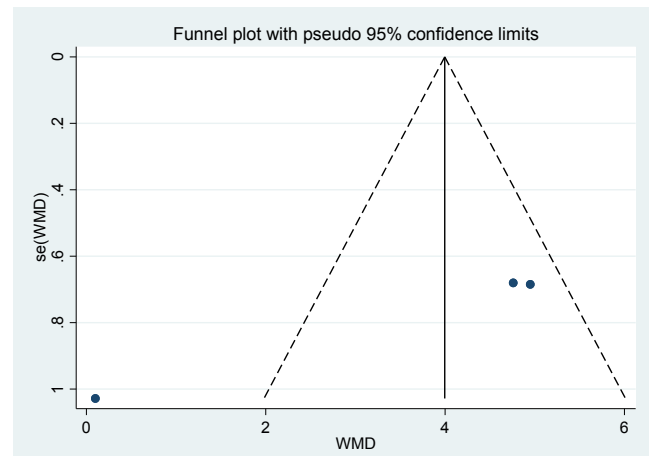
S68.   Funnel plot of SF-36 PCS (JAK inhibitor vs MTX)
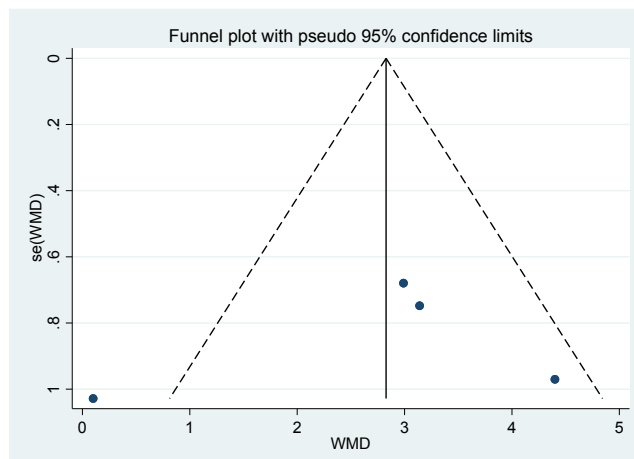
S69.   Funnel plot of FACIT - F (JAK inhibitor vs MTX)

(1) JAK inhibitors compared to placebo

studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
ORAL Solo 2012	1.22 (0.76, 1.96)	126/489	27/122	2.86
Kivitz 2016	1.39 (0.83, 2.33)	161/306	32/72	2.71
Vanhouette 2017 II	1.43 (0.49, 4.16)	37/74	7/17	1.33
FINCH3 2019	1.68 (1.26, 2.25)	503/623	297/416	3.46
Genovese 2017	1.68 (0.87, 3.24)	98/238	15/51	2.27
RA-BUILD 2017	1.83 (1.23, 2.73)	128/456	40/228	3.11
RA-BEGIN 2017	2.20 (1.44, 3.37)	168/215	130/210	3.01
DARWIN1 2016	2.24 (1.41, 3.56)	325/508	38/86	2.89
Kremer 2012	2.25 (1.32, 3.84)	232/438	23/69	2.65
Luchi 2013	2.44 (0.49, 12.01)	19/32	3/8	0.73
ORAL Step 2013	2.52 (1.58, 4.02)	119/265	32/131	2.88
Greenwald 2010	2.55 (1.08, 6.01)	51/93	10/31	1.74
Conaghan 2016	2.57 (0.92, 7.15)	27/35	21/37	1.41
BALANCE II 2016	2.64 (1.40, 4.98)	172/246	22/47	2.33
ORAL Sync 2013	2.67 (1.84, 3.87)	344/633	49/159	3.20
FINCH1 2019	2.74 (2.18, 3.45)	699/955	237/475	3.63
ORAL Standard 2012	2.75 (1.72, 4.38)	204/392	30/106	2.88
I4V-MC-JADA 2015	2.82 (1.71, 4.63)	134/203	40/98	2.78
RA-BEACON 2016	2.90 (1.96, 4.30)	183/351	48/176	3.13
BALANCE I 2016	3.12 (1.68, 5.80)	135/217	19/55	2.38
Fleischmann 2015	3.29 (1.57, 6.91)	94/163	12/41	2.03
SELECT-NEXT 2018	3.34 (2.38, 4.68)	286/440	79/221	3.31
RA-BEAM 2017	3.41 (2.62, 4.45)	339/487	196/488	3.54
FINCH2 2019	3.57 (2.35, 5.42)	185/300	46/148	3.04
RA-BALANCE 2020	3.59 (2.20, 5.86)	85/145	41/145	2.80
SELECT-BEYOND 2018	3.86 (2.58, 5.76)	199/329	48/169	3.10
Fleischmann 2012	4.35 (2.25, 8.42)	150/272	13/59	2.26
SELECT-COMPARE 2019	4.36 (3.45, 5.50)	462/651	234/651	3.63
RAJ3 2019	4.39 (2.63, 7.32)	136/206	31/101	2.73
Robinson 2020	4.50 (1.62, 12.51)	30/42	10/28	1.41
Genovese 2016 II	4.75 (1.07, 21.14)	19/31	3/12	0.81
DARWIN2 2017	5.22 (2.91, 9.37)	144/211	21/72	2.49
SELECT-SUNRISE 2020	5.24 (2.62, 10.49)	118/148	21/49	2.16
RAJ4 2019	5.74 (3.76, 8.77)	214/348	37/170	3.02
Takeuchi 2016	6.55 (2.70, 15.89)	99/225	6/56	1.68
Genovese 2016	6.58 (3.45, 12.55)	171/287	13/71	2.30
Kremer 2009	7.83 (4.18, 14.65)	152/199	19/65	2.36
Tanaka 2016	8.10 (3.72, 17.60)	75/96	15/49	1.94
Vanhouette 2017 I	10.00 (2.00, 50.04)	20/24	4/12	0.72
Tanaka 2015	13.43 (6.04, 29.84)	188/265	8/52	1.88
Tanaka 2011	23.45 (7.37, 74.63)	86/108	4/28	1.19
Boyle 2015	42.38 (2.13, 843.28)	9/15	0/14	0.24
Overall (I-squared = 71.2%, p = 0.000)	3.31 (2.85, 3.86)	7126/11761	1981/5295	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S70. Forest plot of studies comparing the number of patients reaching at least 20% in ACR criteria between patients treated with JAK inhibitors and placebo within 6 months (ACR20 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
Kivitz 2016	1.26 (0.71, 2.24)	95/306	19/72	3.41
FINCH3 2019	1.79 (1.39, 2.30)	374/623	190/416	5.21
Conaghan 2016	2.23 (0.85, 5.87)	17/35	11/37	1.92
RA-BEGIN 2017	2.25 (1.53, 3.32)	136/215	91/210	4.44
Kremer 2012	2.28 (1.18, 4.38)	142/438	12/69	3.04
Luchi 2013	2.65 (0.46, 15.15)	15/32	2/8	0.76
BALANCE I 2016	2.75 (1.28, 5.93)	76/217	9/55	2.57
FINCH1 2019	2.90 (2.23, 3.76)	398/955	94/475	5.16
Genovese 2017	2.90 (1.10, 7.64)	57/238	5/51	1.92
RA-BUILD 2017	2.90 (1.57, 5.37)	68/456	13/228	3.22
Greenwald 2010	3.21 (1.03, 10.02)	30/93	4/31	1.53
FINCH2 2019	3.41 (2.05, 5.68)	112/300	22/148	3.76
DARWIN1 2016	3.50 (1.89, 6.48)	195/508	13/86	3.22
Vanhouette 2017 I	3.67 (0.39, 34.65)	6/24	1/12	0.49
BALANCE II 2016	3.69 (1.66, 8.23)	106/246	8/47	2.44
RA-BEACON 2016	3.70 (2.03, 6.72)	85/351	14/176	3.31
I4V-MC-JADA 2015	3.87 (1.89, 7.94)	62/203	10/98	2.76
SELECT-NEXT 2018	3.87 (2.55, 5.87)	178/440	33/221	4.29
SELECT-BEYOND 2018	4.00 (2.38, 6.73)	115/329	20/169	3.71
RA-BEAM 2017	4.05 (3.01, 5.44)	219/487	82/488	4.97
ORAL Step 2013	4.07 (2.07, 7.99)	72/265	11/131	2.95
Takeuchi 2016	4.17 (1.25, 13.99)	43/225	3/56	1.39
Fleischmann 2012	4.51 (1.87, 10.89)	92/272	6/59	2.18
SELECT-COMPARE 2019	4.62 (3.54, 6.02)	293/651	98/651	5.14
Robinson 2020	4.63 (1.20, 17.92)	15/42	3/28	1.17
Vanhouette 2017 II	4.77 (0.59, 38.64)	17/74	1/17	0.55
RA-BALANCE 2020	4.83 (2.42, 9.61)	44/145	12/145	2.89
DARWIN2 2017	4.98 (2.27, 10.93)	81/211	8/72	2.50
RAJ3 2019	5.85 (2.79, 12.28)	75/206	9/101	2.67
Tanaka 2011	6.00 (1.95, 18.46)	54/108	4/28	1.56
SELECT-SUNRISE 2020	6.20 (2.72, 14.12)	81/148	8/49	2.36
Fleischmann 2015	6.45 (1.91, 21.84)	55/163	3/41	1.38
Genovese 2016	6.74 (2.63, 17.28)	97/287	5/71	2.00
Genovese 2016 II	6.95 (0.79, 60.90)	12/31	1/12	0.52
RAJ4 2019	7.38 (4.03, 13.52)	132/348	13/170	3.27
Tanaka 2015	10.08 (3.53, 28.76)	121/265	4/52	1.72
Tanaka 2016	10.79 (3.60, 32.36)	47/96	4/49	1.61
Kremer 2009	13.11 (4.59, 37.44)	92/199	4/65	1.72
Boyle 2015	19.84 (1.00, 394.78)	6/15	0/14	0.28
Overall (I-squared = 56.6%, p = 0.000)	3.79 (3.22, 4.46)	3915/10247	850/4908	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S71. Forest plot of studies comparing the number of patients reaching at least 50% in ACR criteria between patients treated with JAK inhibitors and placebo within 6 months (ACR50 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
Kivitz 2016	1.27 (0.57, 2.84)	42/306	8/72	3.89
Genovese 2017	1.76 (0.59, 5.23)	31/238	4/51	2.65
FINCH3 2019	2.11 (1.61, 2.77)	265/623	108/416	7.69
Genovese 2016 II	2.12 (0.22, 20.27)	5/31	1/12	0.80
Conaghan 2016	2.22 (0.66, 7.43)	9/35	5/37	2.28
RA-BEGIN 2017	2.40 (1.56, 3.68)	85/215	45/210	6.47
Tanaka 2016	2.52 (0.96, 6.64)	25/96	6/49	3.11
Luchi 2013	2.74 (0.29, 25.54)	9/32	1/8	0.82
DARWIN1 2016	2.98 (1.33, 6.64)	106/508	7/86	3.90
Boyle 2015	3.00 (0.11, 79.91)	1/15	0/14	0.40
SELECT-BEYOND 2018	3.01 (1.53, 5.91)	57/329	11/169	4.65
FINCH2 2019	3.03 (1.49, 6.14)	54/300	10/148	4.45
Kremer 2012	3.36 (1.19, 9.50)	75/438	4/69	2.83
FINCH1 2019	4.00 (2.71, 5.90)	214/955	32/475	6.78
Greenwald 2010	4.02 (0.50, 32.52)	11/93	1/31	0.92
BALANCE II 2016	4.42 (1.32, 14.78)	57/246	3/47	2.29
Takeuchi 2016	4.50 (0.59, 34.52)	17/225	1/56	0.96
RA-BEAM 2017	4.71 (2.92, 7.58)	92/487	23/488	6.09
SELECT-NEXT 2018	5.01 (2.75, 9.15)	105/440	13/221	5.15
Genovese 2016	5.27 (1.24, 22.37)	38/287	2/71	1.73
DARWIN2 2017	5.36 (1.24, 23.08)	28/211	2/72	1.70
Fleischmann 2012	5.65 (1.33, 23.99)	45/272	2/59	1.73
RA-BEACON 2016	5.84 (2.06, 16.58)	42/351	4/176	2.82
SELECT-COMPARE 2019	6.26 (4.22, 9.26)	163/651	33/651	6.76
RA-BUILD 2017	6.43 (1.96, 21.11)	36/456	3/228	2.34
Fleischmann 2015	6.91 (0.91, 52.64)	24/163	1/41	0.97
Tanaka 2011	6.91 (0.89, 53.66)	22/108	1/28	0.95
BALANCE I 2016	6.93 (1.63, 29.54)	45/217	2/55	1.72
RA-BALANCE 2020	7.64 (1.70, 34.26)	14/145	2/145	1.63
Robinson 2020	8.36 (0.44, 157.46)	5/42	0/28	0.49
Kremer 2009	8.43 (1.98, 35.86)	42/199	2/65	1.72
ORAL Step 2013	8.86 (2.09, 37.56)	32/265	2/131	1.73
I4V-MC-JADA 2015	8.98 (2.11, 38.30)	32/203	2/98	1.72
RAJ4 2019	9.00 (3.21, 25.18)	62/348	4/170	2.87
SELECT-SUNRISE 2020	18.39 (2.46, 137.64)	41/148	1/49	0.99
Tanaka 2015	21.66 (2.94, 159.50)	79/265	1/52	1.00
RAJ3 2019	25.61 (3.47, 188.98)	42/206	1/101	1.00
Overall (I-squared = 43.3%, p = 0.003)	4.05 (3.28, 5.01)	2052/10149	348/4879	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S72. Forest plot of studies comparing the number of patients reaching at least 70% in ACR criteria between patients treated with JAK inhibitors and placebo within 6 months (ACR70 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
Conaghan 2016	0.65 (0.21, 1.99)	25/34	30/37	2.24
RA-BUILD 2017	1.33 (0.92, 1.91)	133/456	54/228	7.98
RA-BALANCE 2020	1.49 (0.92, 2.43)	102/145	89/145	6.44
RA-BEGIN 2017	1.61 (1.04, 2.50)	169/215	146/210	7.04
BALANCE II 2016	1.68 (0.87, 3.26)	184/246	30/47	4.72
BALANCE I 2016	1.79 (0.98, 3.25)	126/217	24/55	5.29
ORAL Step 2013	1.91 (1.25, 2.92)	149/264	53/131	7.21
RA-BEAM 2017	2.12 (1.62, 2.78)	363/487	283/488	9.26
I4V-MC-JADA 2015	2.20 (1.34, 3.60)	116/203	37/98	6.38
RA-BEACON 2016	2.26 (1.57, 3.27)	222/351	76/176	7.95
FINCH2 2019	2.47 (1.64, 3.72)	194/292	64/144	7.42
SELECT-NEXT 2018	2.56 (1.83, 3.59)	304/426	109/221	8.37
Kremer 2009	2.61 (1.34, 5.12)	130/166	29/50	4.62
Fleischmann 2015	2.77 (1.32, 5.80)	87/163	12/41	4.10
Tanaka 2011	4.33 (1.67, 11.23)	65/95	8/24	2.88
Tanaka 2016	5.00 (2.36, 10.60)	64/96	14/49	4.02
Tanaka 2015	7.98 (3.79, 16.80)	170/251	10/48	4.07
Overall (I-squared = 57.0%, p = 0.002)	2.19 (1.81, 2.64)	2603/4107	1068/2192	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S73. Forest plot of studies comparing the number of patients reaching at least the minimum clinically important difference (≥0.22 MCID) in HAQ-DI between patients treated with JAK inhibitors and placebo within 6 months (HAQ-DI improvement)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
BALANCE I 2016	1.81 (0.61, 5.41)	27/217	4/55	4.93
RA-BEACON 2016	2.05 (0.68, 6.24)	16/351	4/176	4.79
DARWIN2 2017	2.11 (0.46, 9.66)	12/211	2/72	2.56
SELECT-BEYOND 2018	2.17 (0.98, 4.82)	32/329	8/169	9.21
RA-BEGIN 2017	2.34 (1.36, 4.01)	48/215	23/210	19.83
BALANCE II 2016	2.66 (0.61, 11.61)	26/246	2/47	2.73
SELECT-NEXT 2018	3.57 (1.58, 8.04)	46/440	7/221	8.90
RA-BEAM 2017	3.99 (2.02, 7.85)	41/487	11/488	12.69
SELECT-COMPARE 2019	4.74 (2.87, 7.81)	85/651	20/651	22.91
DARWIN1 2016	4.89 (1.17, 20.46)	53/508	2/86	2.90
RA-BUILD 2017	4.91 (1.13, 21.28)	19/456	2/228	2.76
RA-BALANCE 2020	5.07 (0.24, 106.53)	2/145	0/145	0.64
SELECT-SUNRISE 2020	8.83 (1.16, 67.22)	23/148	1/49	1.44
I4V-MC-JADA 2015	11.19 (1.48, 84.47)	21/203	1/98	1.46
RAJ4 2019	17.88 (2.42, 131.92)	33/343	1/169	1.49
RAJ3 2019	21.13 (1.26, 353.63)	19/204	0/100	0.75
Overall (I-squared = 1.0%, p = 0.441)	3.41 (2.67, 4.36)	503/5154	88/2964	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S74. Forest plot of studies comparing the number of patients reaching remission defined by CDAI (≤ 2.8) between patients treated with JAK inhibitors and placebo within 6 months (CDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S75. Forest plot of studies comparing the number of patients reaching remission defined by SDAI (≤ 3.3) between patients treated with JAK inhibitors and placebo within 6 months (SDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
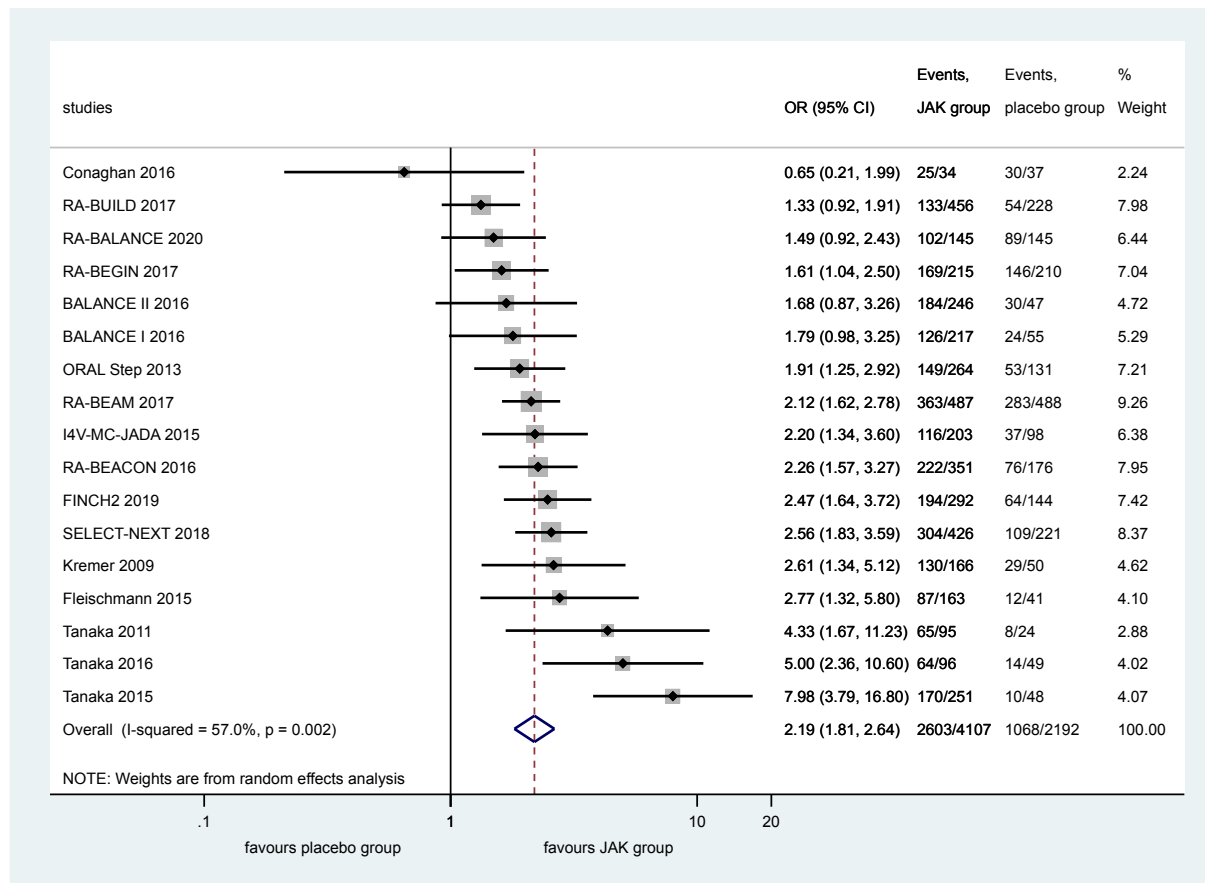
studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
Greenwald 2010	1.42 (0.49, 4.18)	20/93	5/31	2.43
Genovese 2017	1.75 (0.65, 4.71)	37/231	5/51	2.75
RA-BEGIN 2017	2.17 (1.43, 3.30)	87/215	50/210	6.55
Tanaka 2016	2.26 (1.03, 4.97)	38/96	11/49	3.72
FINCH3 2019	2.46 (1.89, 3.20)	313/623	121/416	8.01
Vanhouette 2017 II	2.50 (0.30, 20.98)	10/74	1/17	0.77
BALANCE I 2016	2.63 (1.13, 6.12)	60/220	7/56	3.39
Takeuchi 2016	2.82 (0.83, 9.59)	31/225	3/56	2.00
DARWIN2 2017	2.88 (1.39, 5.97)	67/211	10/72	4.06
BALANCE II 2016	3.13 (1.35, 7.25)	84/249	7/50	3.42
RA-BUILD 2017	3.13 (1.51, 6.48)	52/456	9/228	4.07
SELECT-BEYOND 2018	3.44 (1.94, 6.08)	87/329	16/169	5.20
DARWIN1 2016	3.52 (1.49, 8.28)	106/508	6/86	3.34
FINCH2 2019	3.58 (1.87, 6.83)	72/300	12/148	4.61
RA-BEACON 2016	3.73 (1.65, 8.44)	47/351	7/176	3.55
SELECT-NEXT 2018	3.79 (2.33, 6.16)	130/440	22/221	5.93
FINCH1 2019	3.96 (2.82, 5.57)	275/955	44/475	7.29
RA-BALANCE 2020	4.68 (1.53, 14.28)	17/145	4/145	2.30
ORAL Solo 2012	5.21 (2.23, 12.19)	104/483	6/120	3.37
SELECT-COMPARE 2019	6.42 (4.46, 9.25)	189/651	39/651	7.06
Robinson 2020	6.66 (0.34, 128.78)	4/42	0/28	0.41
I4V-MC-JADA 2015	6.69 (2.33, 19.20)	45/203	4/98	2.51
Vanhouette 2017 I	7.05 (0.36, 138.95)	5/24	0/12	0.41
RA-BEAM 2017	7.19 (4.43, 11.66)	119/487	21/488	5.95
Luchi 2013	7.93 (0.42, 150.78)	10/32	0/8	0.42
RAJ3 2019	7.97 (3.09, 20.58)	60/203	5/100	2.91
Genovese 2016	11.34 (2.71, 47.44)	71/287	2/71	1.55
SELECT-SUNRISE 2020	14.14 (4.21, 47.49)	71/148	3/49	2.03
Overall (I-squared = 48.7%, p = 0.002)	3.75 (3.08, 4.56)	2211/8281	420/4281	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S76. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-CRP (<2.6) between patients treated with JAK inhibitors and placebo within 6 months (DAS28-CRP remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
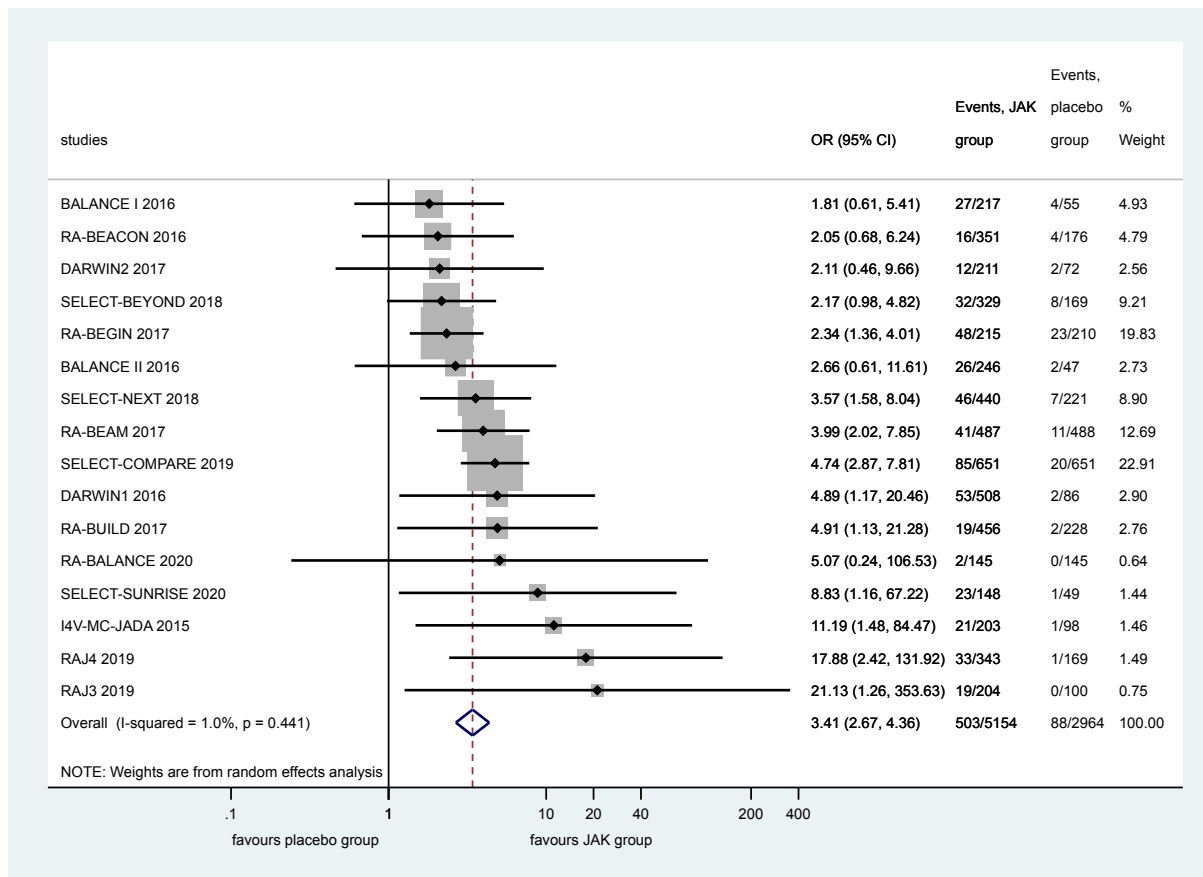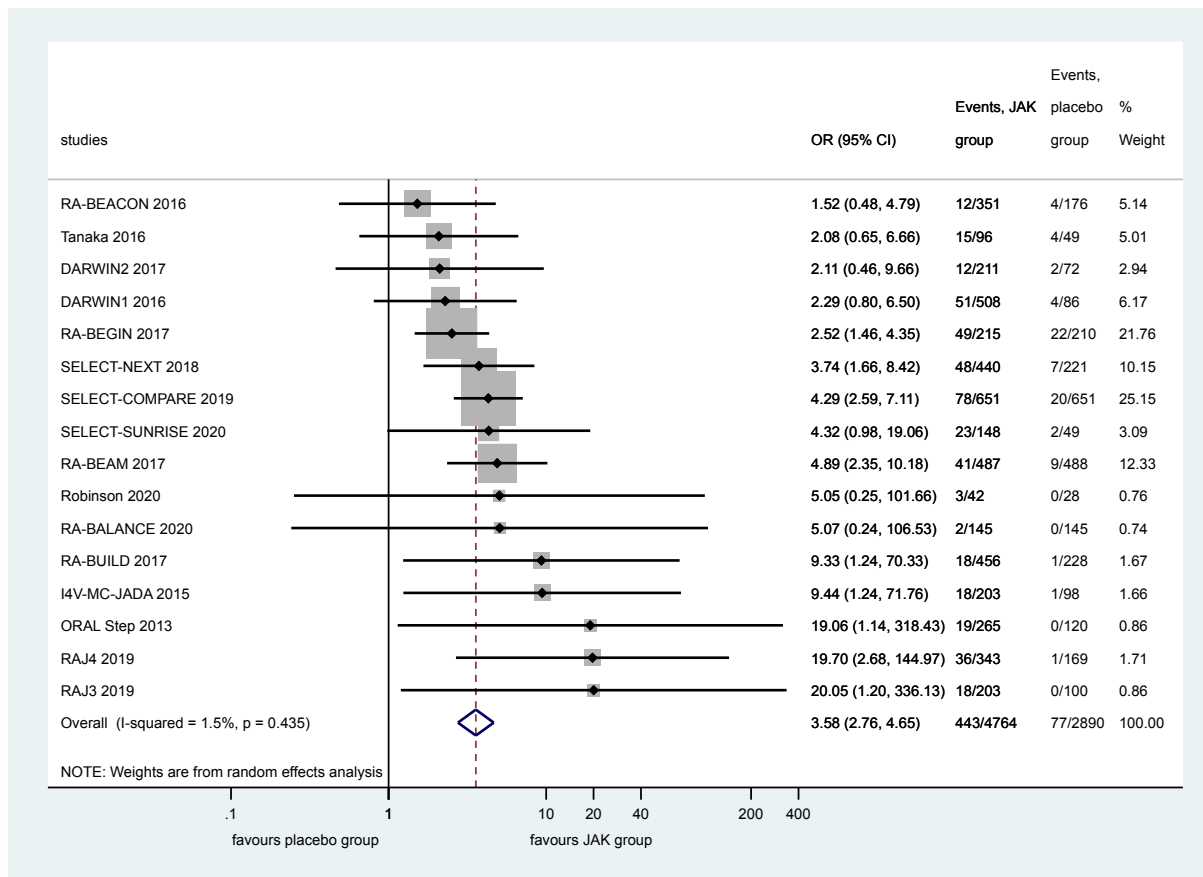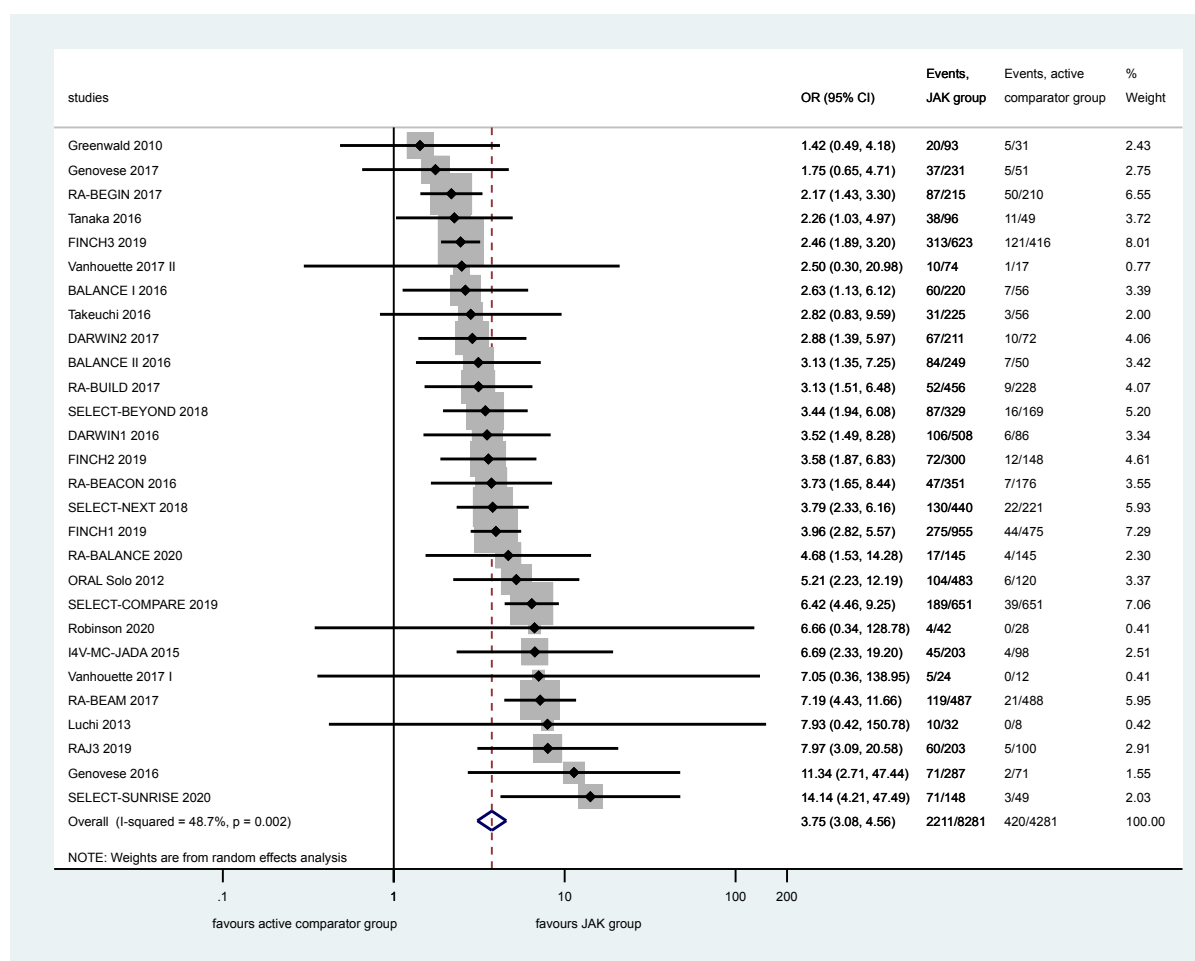
studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
ORAL Solo 2012	1.68 (0.64, 4.41)	33/461	5/114	9.19
Conaghan 2016	1.97 (0.57, 6.75)	8/34	5/37	6.24
Takeuchi 2016	2.21 (0.49, 9.84)	17/225	2/56	4.47
RA-BEGIN 2017	2.42 (1.44, 4.07)	53/215	25/210	19.61
Genovese 2017	3.52 (0.81, 15.24)	29/231	2/51	4.63
Tanaka 2016	3.68 (0.80, 17.02)	13/96	2/49	4.29
ORAL Sync 2013	4.37 (1.56, 12.23)	60/571	4/153	8.33
Robinson 2020	5.05 (0.25, 101.66)	3/42	0/28	1.22
RA-BEAM 2017	5.07 (2.61, 9.86)	51/487	11/488	15.15
RA-BUILD 2017	5.18 (1.20, 22.37)	20/456	2/228	4.65
RA-BEACON 2016	5.54 (1.28, 23.89)	21/351	2/176	4.65
RAJ4 2019	8.36 (2.98, 23.46)	55/343	4/179	8.28
ORAL Standard 2012	9.38 (1.27, 69.55)	33/353	1/92	2.64
I4V-MC-JADA 2015	14.88 (1.99, 111.20)	27/203	1/98	2.62
RAJ3 2019	17.07 (2.29, 127.08)	30/204	1/100	2.62
Tanaka 2015	24.61 (1.49, 405.35)	50/265	0/52	1.40
Overall (I-squared = 18.6%, p = 0.241)	3.99 (2.84, 5.59)	503/4537	67/2111	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S77. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-ESR (<2.6) between patients treated with JAK inhibitors and placebo within 6 months (DAS28-ESR remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
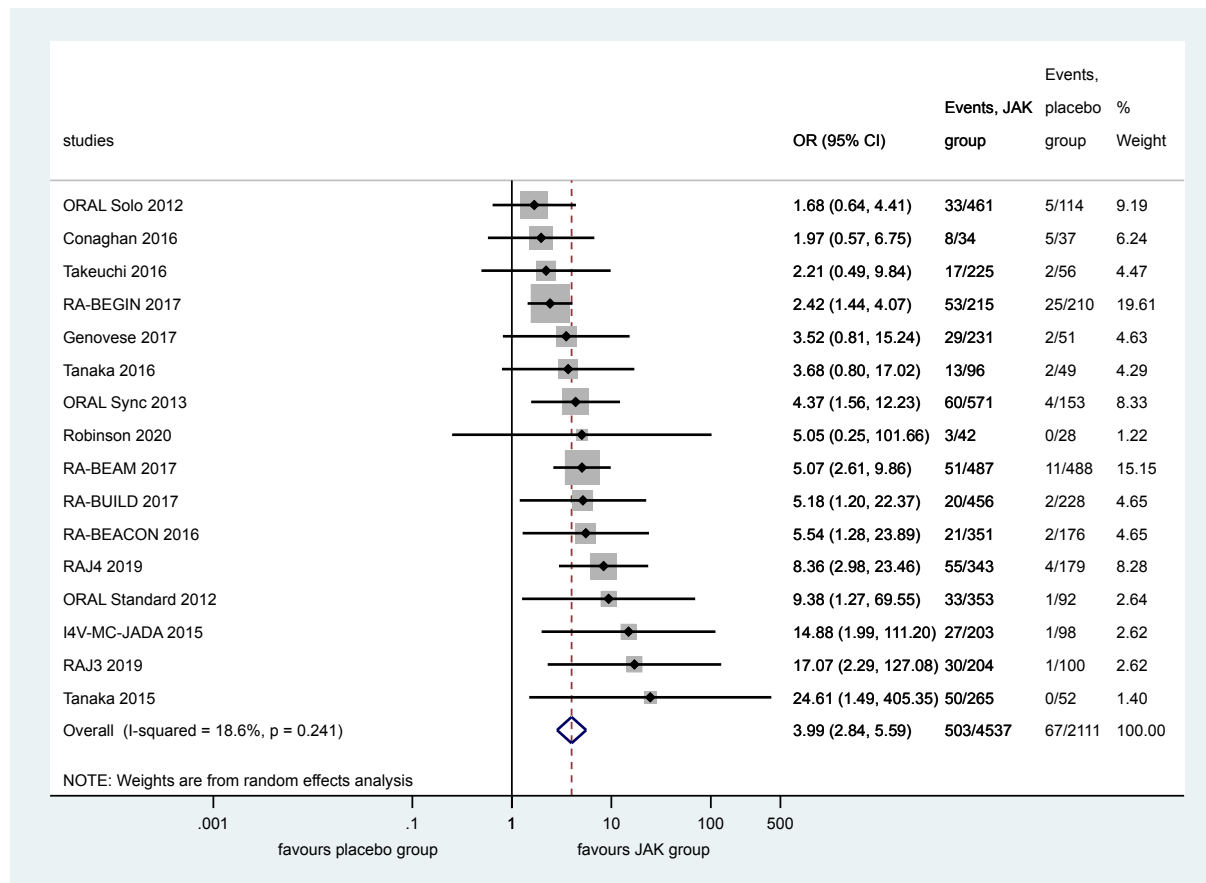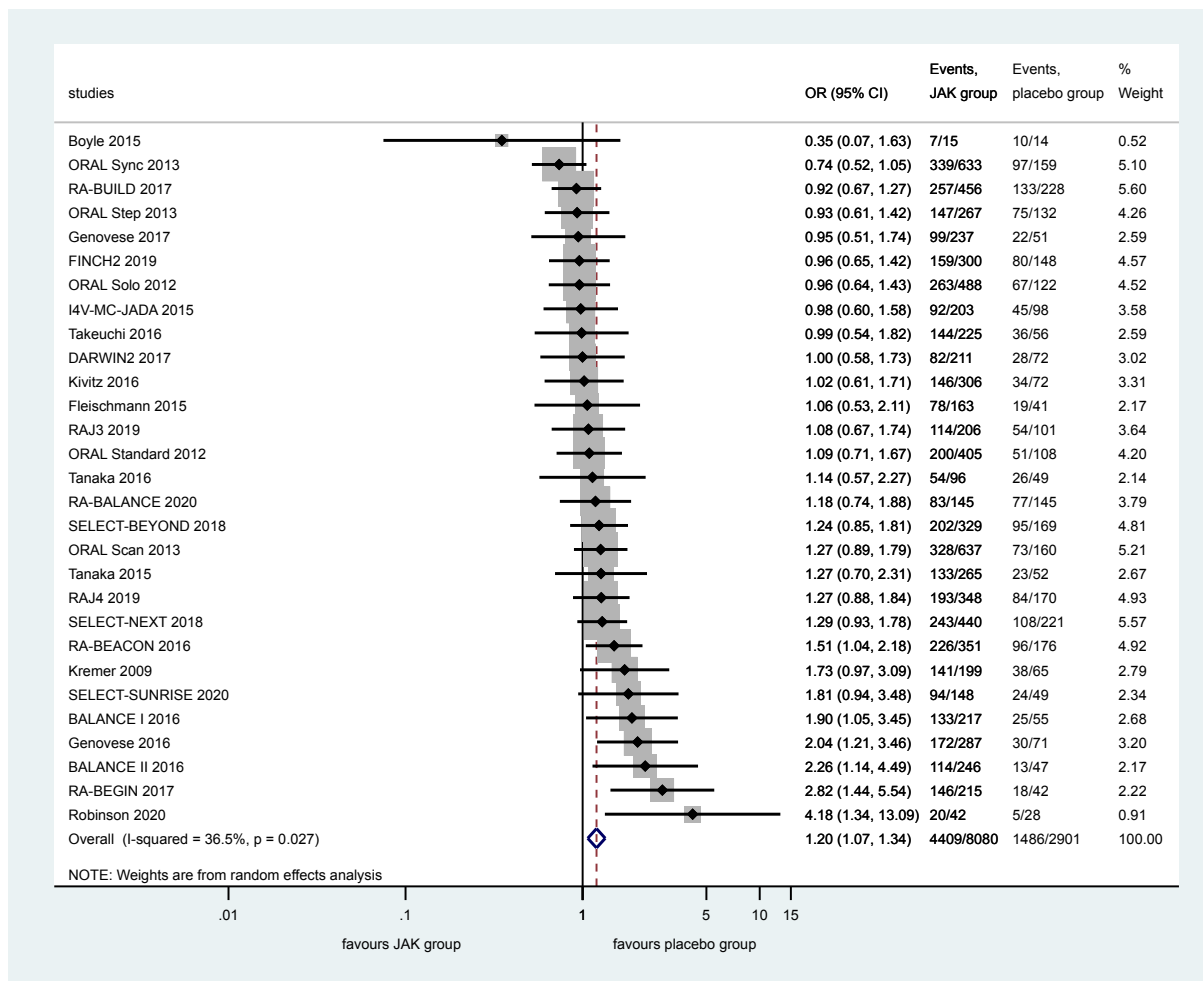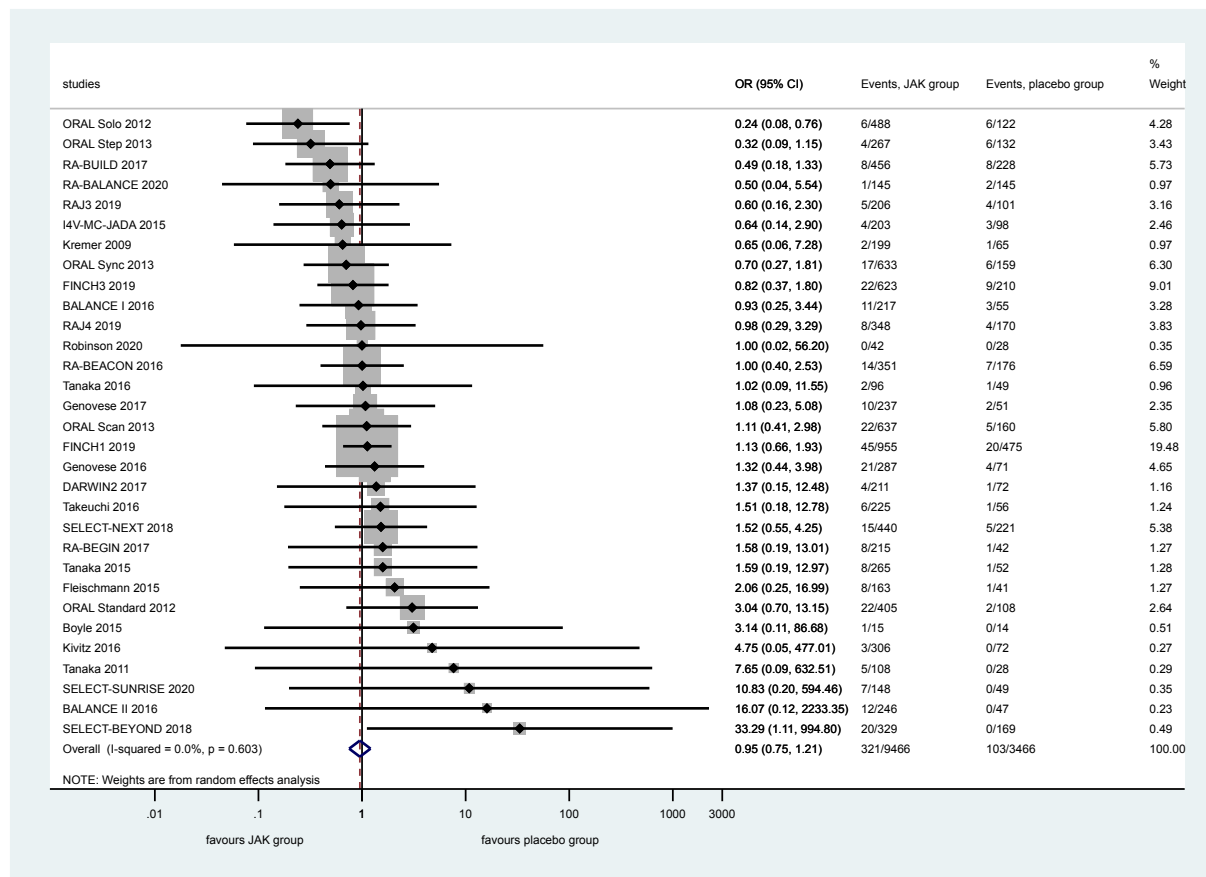
studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
Boyle 2015	0.35 (0.07, 1.63)	7/15	10/14	0.52
ORAL Sync 2013	0.74 (0.52, 1.05)	339/633	97/159	5.10
RA-BUILD 2017	0.92 (0.67, 1.27)	257/456	133/228	5.60
ORAL Step 2013	0.93 (0.61, 1.42)	147/267	75/132	4.26
Genovese 2017	0.95 (0.51, 1.74)	99/237	22/51	2.59
FINCH2 2019	0.96 (0.65, 1.42)	159/300	80/148	4.57
ORAL Solo 2012	0.96 (0.64, 1.43)	263/488	67/122	4.52
I4V-MC-JADA 2015	0.98 (0.60, 1.58)	92/203	45/98	3.58
Takeuchi 2016	0.99 (0.54, 1.82)	144/225	36/56	2.59
DARWIN2 2017	1.00 (0.58, 1.73)	82/211	28/72	3.02
Kivitz 2016	1.02 (0.61, 1.71)	146/306	34/72	3.31
Fleischmann 2015	1.06 (0.53, 2.11)	78/163	19/41	2.17
RAJ3 2019	1.08 (0.67, 1.74)	114/206	54/101	3.64
ORAL Standard 2012	1.09 (0.71, 1.67)	200/405	51/108	4.20
Tanaka 2016	1.14 (0.57, 2.27)	54/96	26/49	2.14
RA-BALANCE 2020	1.18 (0.74, 1.88)	83/145	77/145	3.79
SELECT-BEYOND 2018	1.24 (0.85, 1.81)	202/329	95/169	4.81
ORAL Scan 2013	1.27 (0.89, 1.79)	328/637	73/160	5.21
Tanaka 2015	1.27 (0.70, 2.31)	133/265	23/52	2.67
RAJ4 2019	1.27 (0.88, 1.84)	193/348	84/170	4.93
SELECT-NEXT 2018	1.29 (0.93, 1.78)	243/440	108/221	5.57
RA-BEACON 2016	1.51 (1.04, 2.18)	226/351	96/176	4.92
Kremer 2009	1.73 (0.97, 3.09)	141/199	38/65	2.79
SELECT-SUNRISE 2020	1.81 (0.94, 3.48)	94/148	24/49	2.34
BALANCE I 2016	1.90 (1.05, 3.45)	133/217	25/55	2.68
Genovese 2016	2.04 (1.21, 3.46)	172/287	30/71	3.20
BALANCE II 2016	2.26 (1.14, 4.49)	114/246	13/47	2.17
RA-BEGIN 2017	2.82 (1.44, 5.54)	146/215	18/42	2.22
Robinson 2020	4.18 (1.34, 13.09)	20/42	5/28	0.91
Overall (I-squared = 36.5%, p = 0.027)	1.20 (1.07, 1.34)	4409/8080	1486/2901	100.00

NOTE: Weights are from random effects analysis

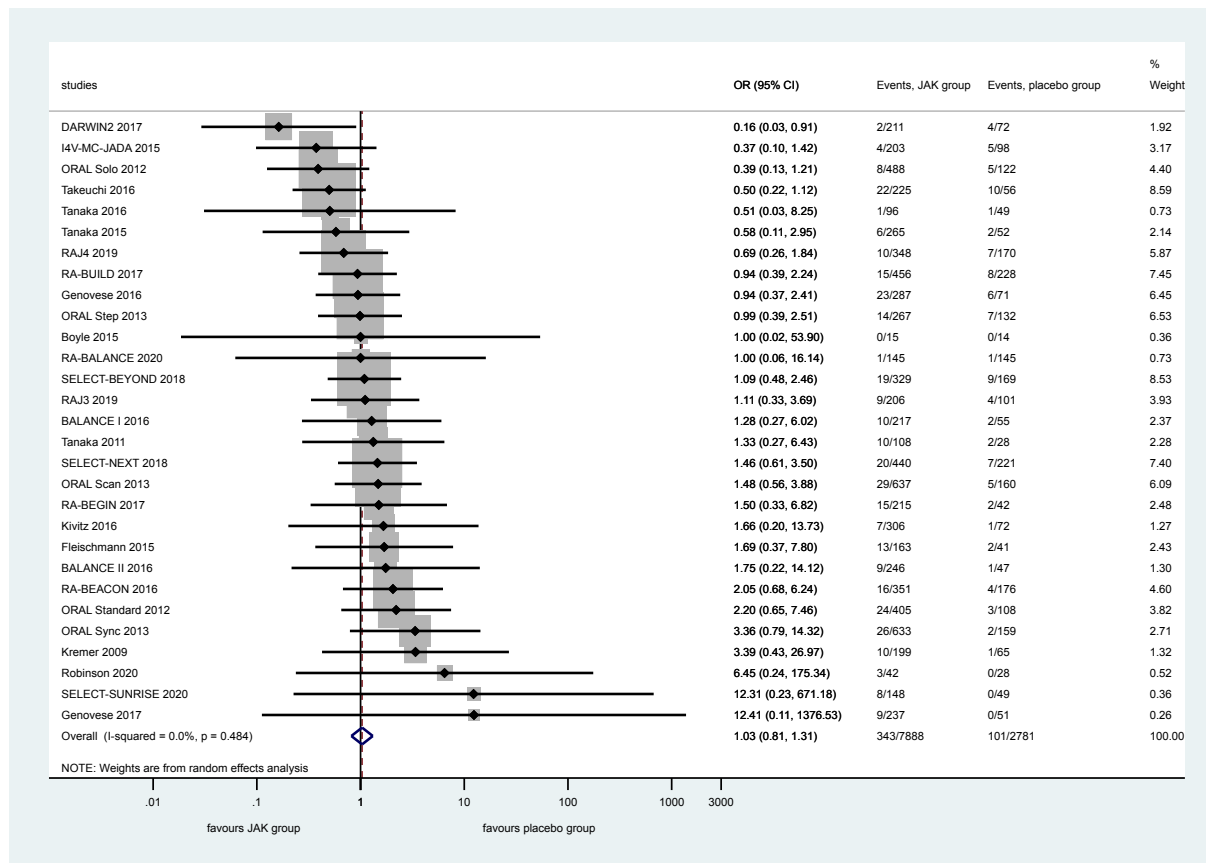favours JAK group   favours placebo group

S78. Forest plot of studies comparing the number of patients experiencing side effects during the study between patients treated with JAK inhibitors and placebo within 6 months (Side effects)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
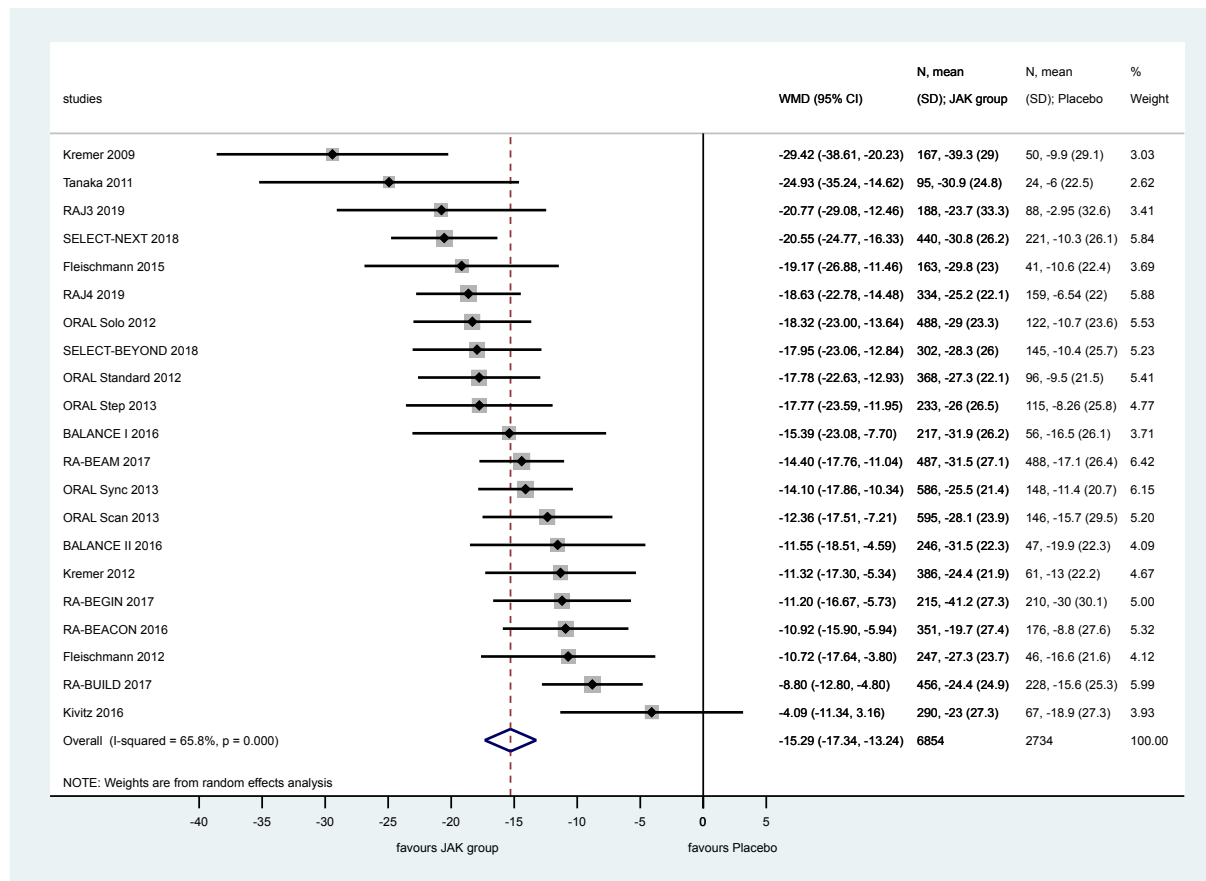
studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
ORAL Solo 2012	0.24 (0.08, 0.76)	6/488	6/122	4.28
ORAL Step 2013	0.32 (0.09, 1.15)	4/267	6/132	3.43
RA-BUILD 2017	0.49 (0.18, 1.33)	8/456	8/228	5.73
RA-BALANCE 2020	0.50 (0.04, 5.54)	1/145	2/145	0.97
RAJ3 2019	0.60 (0.16, 2.30)	5/206	4/101	3.16
I4V-MC-JADA 2015	0.64 (0.14, 2.90)	4/203	3/98	2.46
Kremer 2009	0.65 (0.06, 7.28)	2/199	1/65	0.97
ORAL Sync 2013	0.70 (0.27, 1.81)	17/633	6/159	6.30
FINCH3 2019	0.82 (0.37, 1.80)	22/623	9/210	9.01
BALANCE I 2016	0.93 (0.25, 3.44)	11/217	3/55	3.28
RAJ4 2019	0.98 (0.29, 3.29)	8/348	4/170	3.83
Robinson 2020	1.00 (0.02, 56.20)	0/42	0/28	0.35
RA-BEACON 2016	1.00 (0.40, 2.53)	14/351	7/176	6.59
Tanaka 2016	1.02 (0.09, 11.55)	2/96	1/49	0.96
Genovese 2017	1.08 (0.23, 5.08)	10/237	2/51	2.35
ORAL Scan 2013	1.11 (0.41, 2.98)	22/637	5/160	5.80
FINCH1 2019	1.13 (0.66, 1.93)	45/955	20/475	19.48
Genovese 2016	1.32 (0.44, 3.98)	21/287	4/71	4.65
DARWIN2 2017	1.37 (0.15, 12.48)	4/211	1/72	1.16
Takeuchi 2016	1.51 (0.18, 12.78)	6/225	1/56	1.24
SELECT-NEXT 2018	1.52 (0.55, 4.25)	15/440	5/221	5.38
RA-BEGIN 2017	1.58 (0.19, 13.01)	8/215	1/42	1.27
Tanaka 2015	1.59 (0.19, 12.97)	8/265	1/52	1.28
Fleischmann 2015	2.06 (0.25, 16.99)	8/163	1/41	1.27
ORAL Standard 2012	3.04 (0.70, 13.15)	22/405	2/108	2.64
Boyle 2015	3.14 (0.11, 86.68)	1/15	0/14	0.51
Kivitz 2016	4.75 (0.05, 477.01)	3/306	0/72	0.27
Tanaka 2011	7.65 (0.09, 632.51)	5/108	0/28	0.29
SELECT-SUNRISE 2020	10.83 (0.20, 594.46)	7/148	0/49	0.35
BALANCE II 2016	16.07 (0.12, 2233.35)	12/246	0/47	0.23
SELECT-BEYOND 2018	33.29 (1.11, 994.80)	20/329	0/169	0.49
Overall (I-squared = 0.0%, p = 0.603)	0.95 (0.75, 1.21)	321/9466	103/3466	100.00

NOTE: Weights are from random effects analysis

S79. Forest plot of studies comparing the number of patients experiencing serious side effects during the study between patients treated with JAK inhibitors and placebo within 6 months (Serious side effects)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
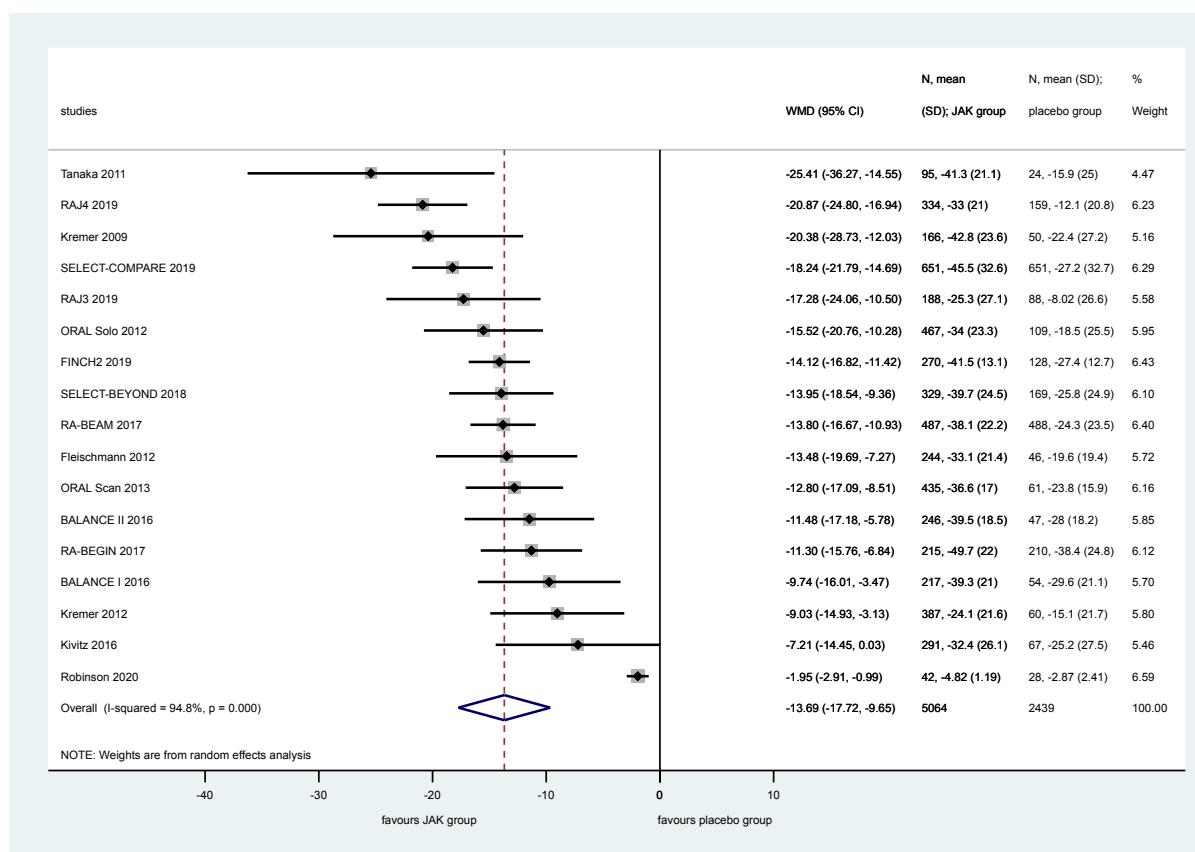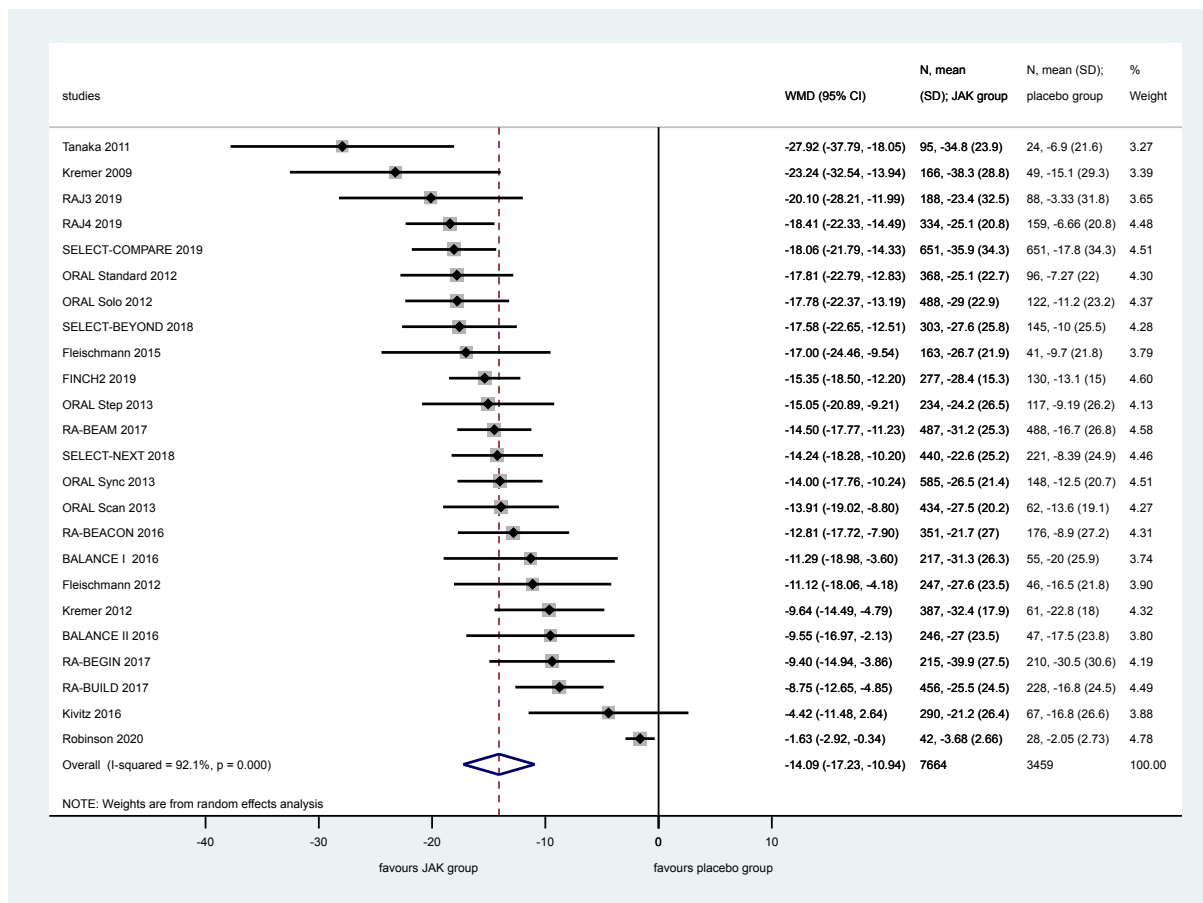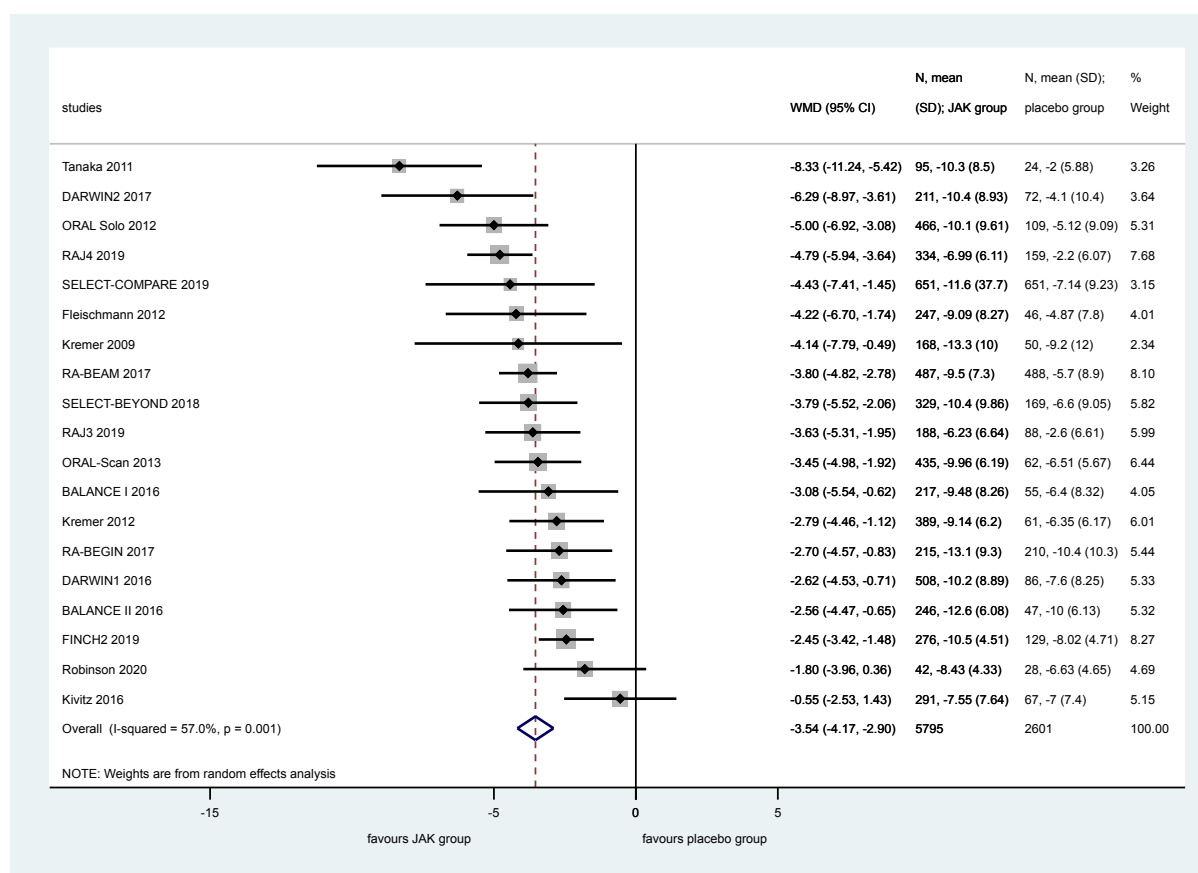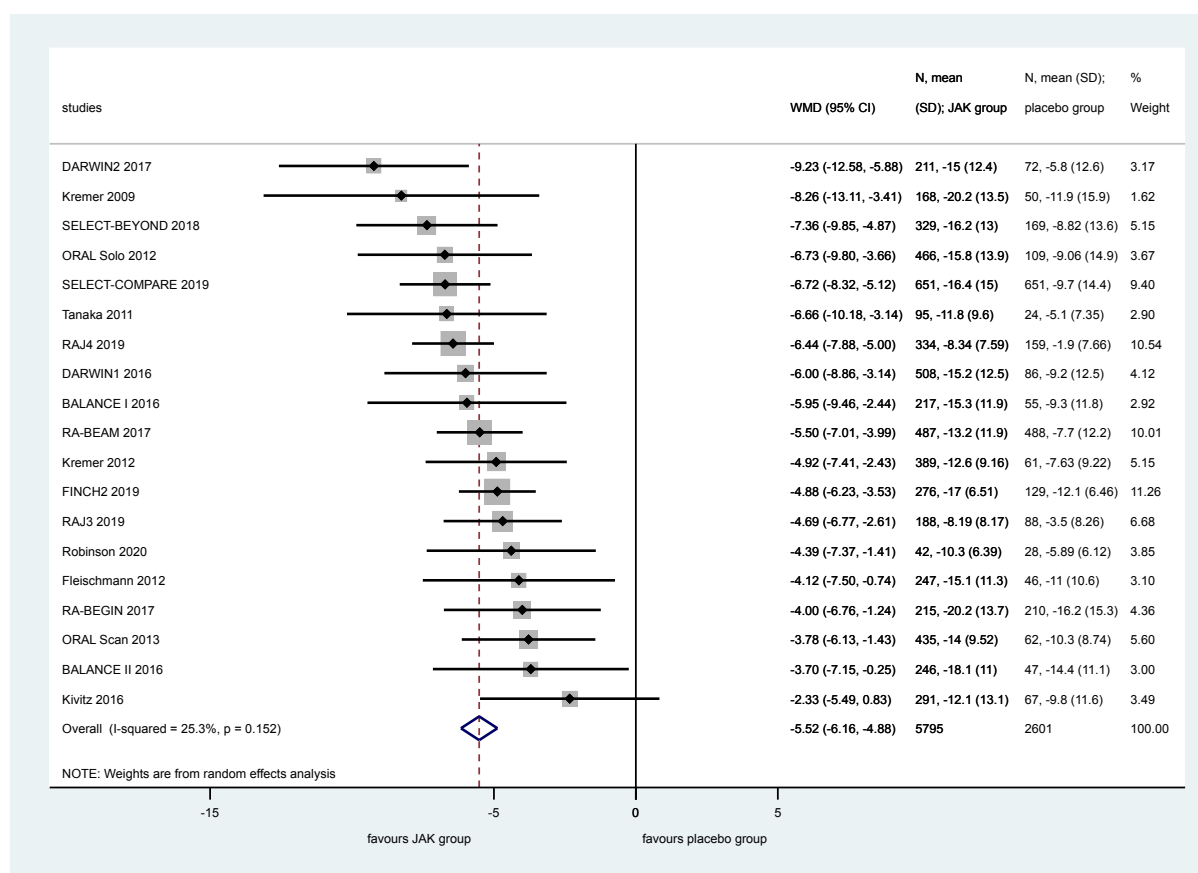
studies	OR (95% CI)	Events, JAK group	Events, placebo group	% Weight
DARWIN2 2017	0.16 (0.03, 0.91)	2/211	4/72	1.92
I4V-MC-JADA 2015	0.37 (0.10, 1.42)	4/203	5/98	3.17
ORAL Solo 2012	0.39 (0.13, 1.21)	8/488	5/122	4.40
Takeuchi 2016	0.50 (0.22, 1.12)	22/225	10/56	8.59
Tanaka 2016	0.51 (0.03, 8.25)	1/96	1/49	0.73
Tanaka 2015	0.58 (0.11, 2.95)	6/265	2/52	2.14
RAJ4 2019	0.69 (0.26, 1.84)	10/348	7/170	5.87
RA-BUILD 2017	0.94 (0.39, 2.24)	15/456	8/228	7.45
Genovese 2016	0.94 (0.37, 2.41)	23/287	6/71	6.45
ORAL Step 2013	0.99 (0.39, 2.51)	14/267	7/132	6.53
Boyle 2015	1.00 (0.02, 53.90)	0/15	0/14	0.36
RA-BALANCE 2020	1.00 (0.06, 16.14)	1/145	1/145	0.73
SELECT-BEYOND 2018	1.09 (0.48, 2.46)	19/329	9/169	8.53
RAJ3 2019	1.11 (0.33, 3.69)	9/206	4/101	3.93
BALANCE I 2016	1.28 (0.27, 6.02)	10/217	2/55	2.37
Tanaka 2011	1.33 (0.27, 6.43)	10/108	2/28	2.28
SELECT-NEXT 2018	1.46 (0.61, 3.50)	20/440	7/221	7.40
ORAL Scan 2013	1.48 (0.56, 3.88)	29/637	5/160	6.09
RA-BEGIN 2017	1.50 (0.33, 6.82)	15/215	2/42	2.48
Kivitz 2016	1.66 (0.20, 13.73)	7/306	1/72	1.27
Fleischmann 2015	1.69 (0.37, 7.80)	13/163	2/41	2.43
BALANCE II 2016	1.75 (0.22, 14.12)	9/246	1/47	1.30
RA-BEACON 2016	2.05 (0.68, 6.24)	16/351	4/176	4.60
ORAL Standard 2012	2.20 (0.65, 7.46)	24/405	3/108	3.82
ORAL Sync 2013	3.36 (0.79, 14.32)	26/633	2/159	2.71
Kremer 2009	3.39 (0.43, 26.97)	10/199	1/65	1.32
Robinson 2020	6.45 (0.24, 175.34)	3/42	0/28	0.52
SELECT-SUNRISE 2020	12.31 (0.23, 671.18)	8/148	0/49	0.36
Genovese 2017	12.41 (0.11, 1376.53)	9/237	0/51	0.26
Overall (I-squared = 0.0%, p = 0.484)	1.03 (0.81, 1.31)	343/7888	101/2781	100.00

NOTE: Weights are from random effects analysis

.01   .1   1   10   100   1000   3000

favours JAK group   favours placebo group

S80. Forest plot of studies comparing the number of patients experiencing symptoms leading to discontinuation of the study drugs between patients treated with JAK inhibitors and placebo within 6 months (Discontinuation)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
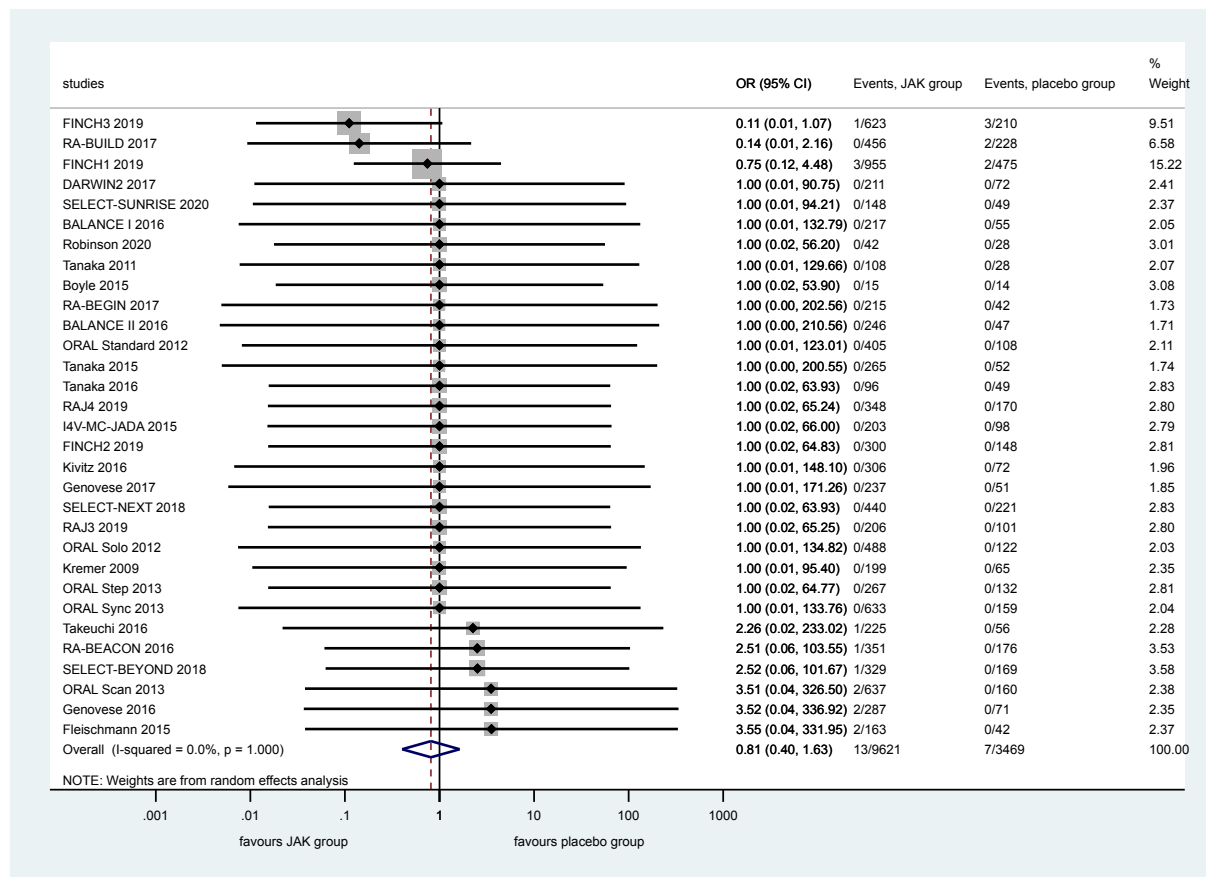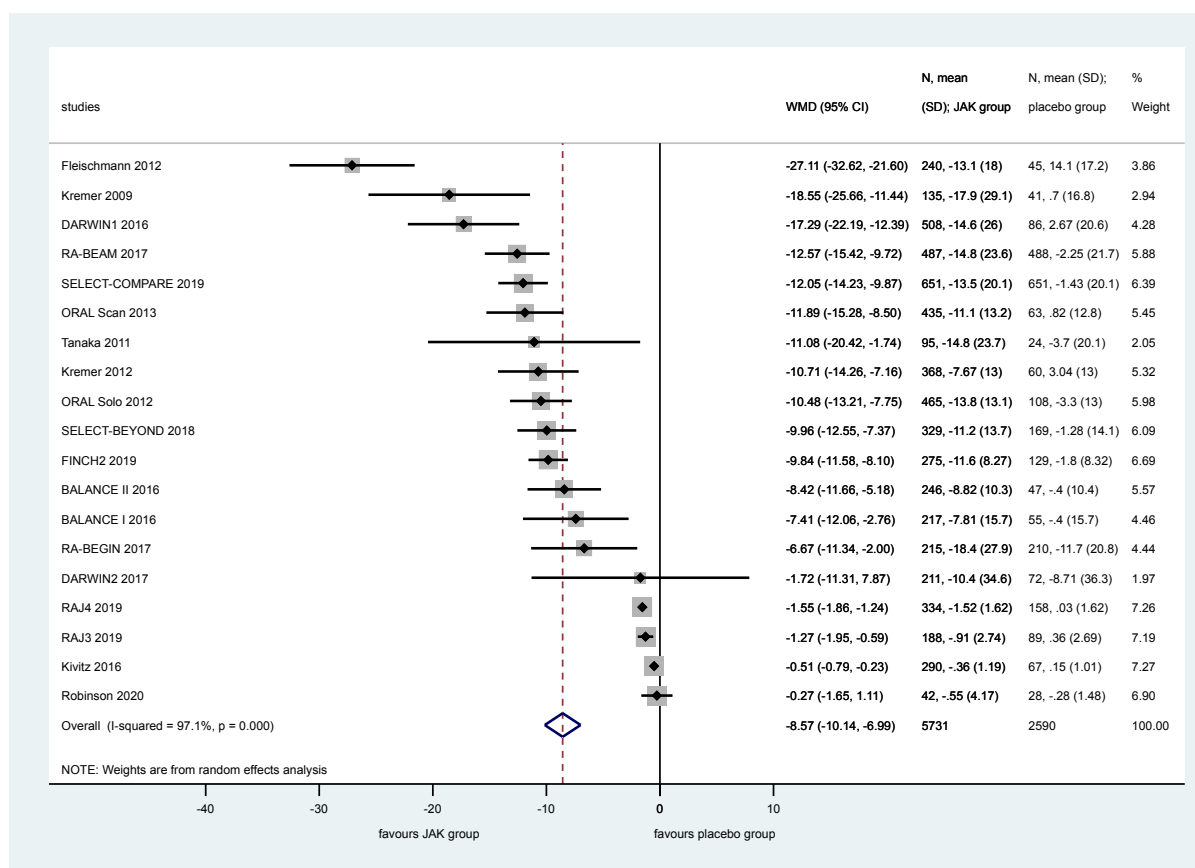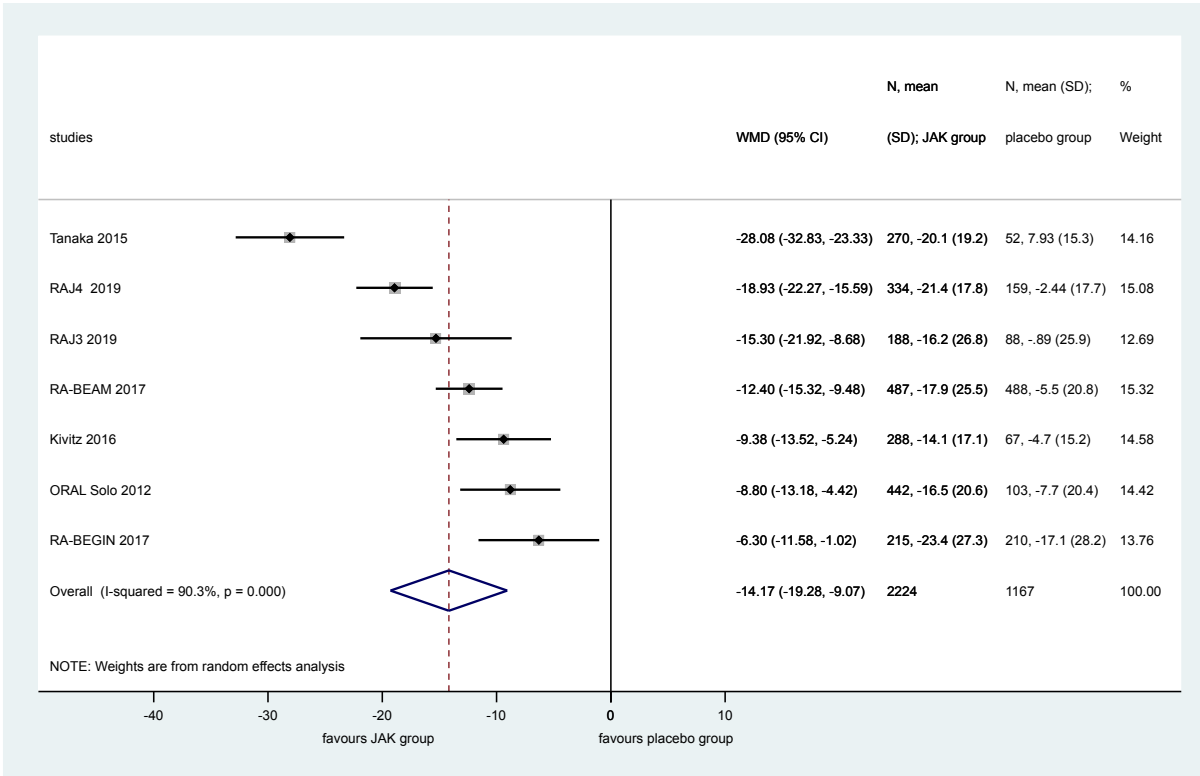
S81. Forest plot of studies comparing the number of patients who dies during study period between patients treated with JAK inhibitors and placebo (Deaths)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); Placebo	% Weight
Kremer 2009	-29.42 (-38.61, -20.23)	167, -39.3 (29)	50, -9.9 (29.1)	3.03
Tanaka 2011	-24.93 (-35.24, -14.62)	95, -30.9 (24.8)	24, -6 (22.5)	2.62
RAJ3 2019	-20.77 (-29.08, -12.46)	188, -23.7 (33.3)	88, -2.95 (32.6)	3.41
SELECT-NEXT 2018	-20.55 (-24.77, -16.33)	440, -30.8 (26.2)	221, -10.3 (26.1)	5.84
Fleischmann 2015	-19.17 (-26.88, -11.46)	163, -29.8 (23)	41, -10.6 (22.4)	3.69
RAJ4 2019	-18.63 (-22.78, -14.48)	334, -25.2 (22.1)	159, -6.54 (22)	5.88
ORAL Solo 2012	-18.32 (-23.00, -13.64)	488, -29 (23.3)	122, -10.7 (23.6)	5.53
SELECT-BEYOND 2018	-17.95 (-23.06, -12.84)	302, -28.3 (26)	145, -10.4 (25.7)	5.23
ORAL Standard 2012	-17.78 (-22.63, -12.93)	368, -27.3 (22.1)	96, -9.5 (21.5)	5.41
ORAL Step 2013	-17.77 (-23.59, -11.95)	233, -26 (26.5)	115, -8.26 (25.8)	4.77
BALANCE I 2016	-15.39 (-23.08, -7.70)	217, -31.9 (26.2)	56, -16.5 (26.1)	3.71
RA-BEAM 2017	-14.40 (-17.76, -11.04)	487, -31.5 (27.1)	488, -17.1 (26.4)	6.42
ORAL Sync 2013	-14.10 (-17.86, -10.34)	586, -25.5 (21.4)	148, -11.4 (20.7)	6.15
ORAL Scan 2013	-12.36 (-17.51, -7.21)	595, -28.1 (23.9)	146, -15.7 (29.5)	5.20
BALANCE II 2016	-11.55 (-18.51, -4.59)	246, -31.5 (22.3)	47, -19.9 (22.3)	4.09
Kremer 2012	-11.32 (-17.30, -5.34)	386, -24.4 (21.9)	61, -13 (22.2)	4.67
RA-BEGIN 2017	-11.20 (-16.67, -5.73)	215, -41.2 (27.3)	210, -30 (30.1)	5.00
RA-BEACON 2016	-10.92 (-15.90, -5.94)	351, -19.7 (27.4)	176, -8.8 (27.6)	5.32
Fleischmann 2012	-10.72 (-17.64, -3.80)	247, -27.3 (23.7)	46, -16.6 (21.6)	4.12
RA-BUILD 2017	-8.80 (-12.80, -4.80)	456, -24.4 (24.9)	228, -15.6 (25.3)	5.99
Kivitz 2016	-4.09 (-11.34, 3.16)	290, -23 (27.3)	67, -18.9 (27.3)	3.93
Overall (I-squared = 65.8%, p = 0.000)	-15.29 (-17.34, -13.24)	6854	2734	100.00

NOTE: Weights are from random effects analysis

S82. Forest plot of studies comparing the LSM change from baseline in pain measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and placebo within 6 months (Pain VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
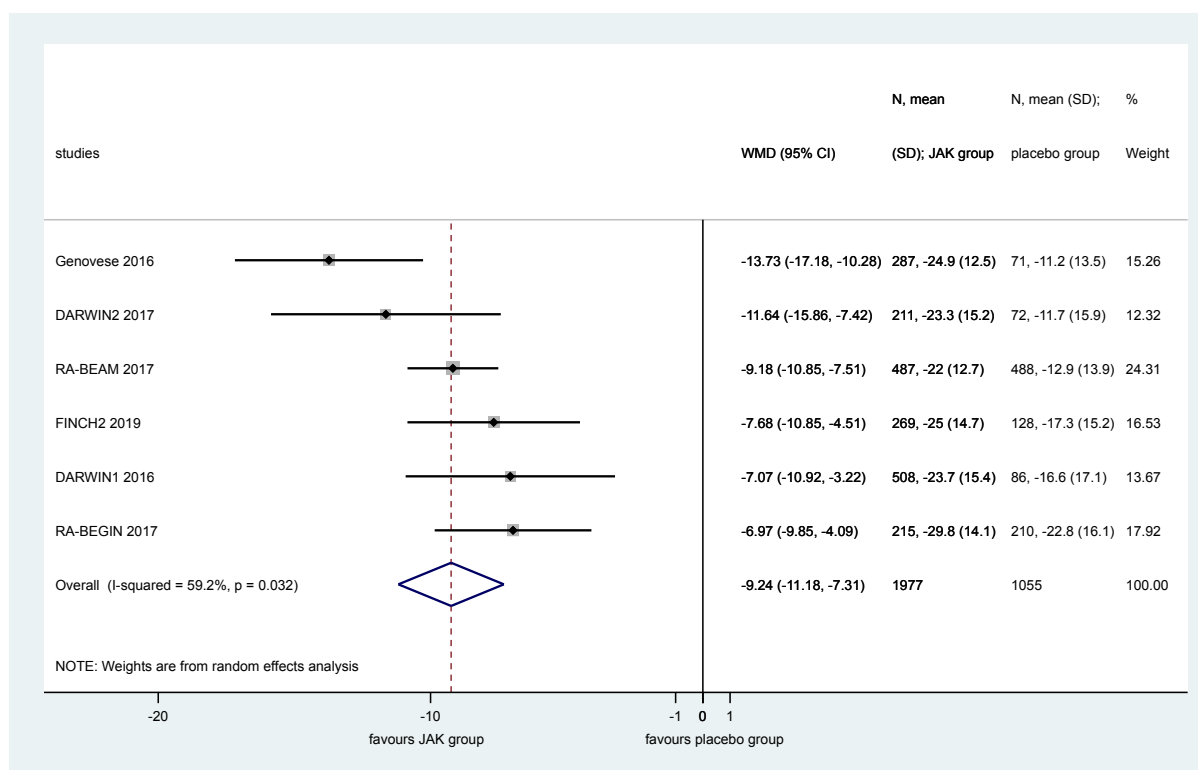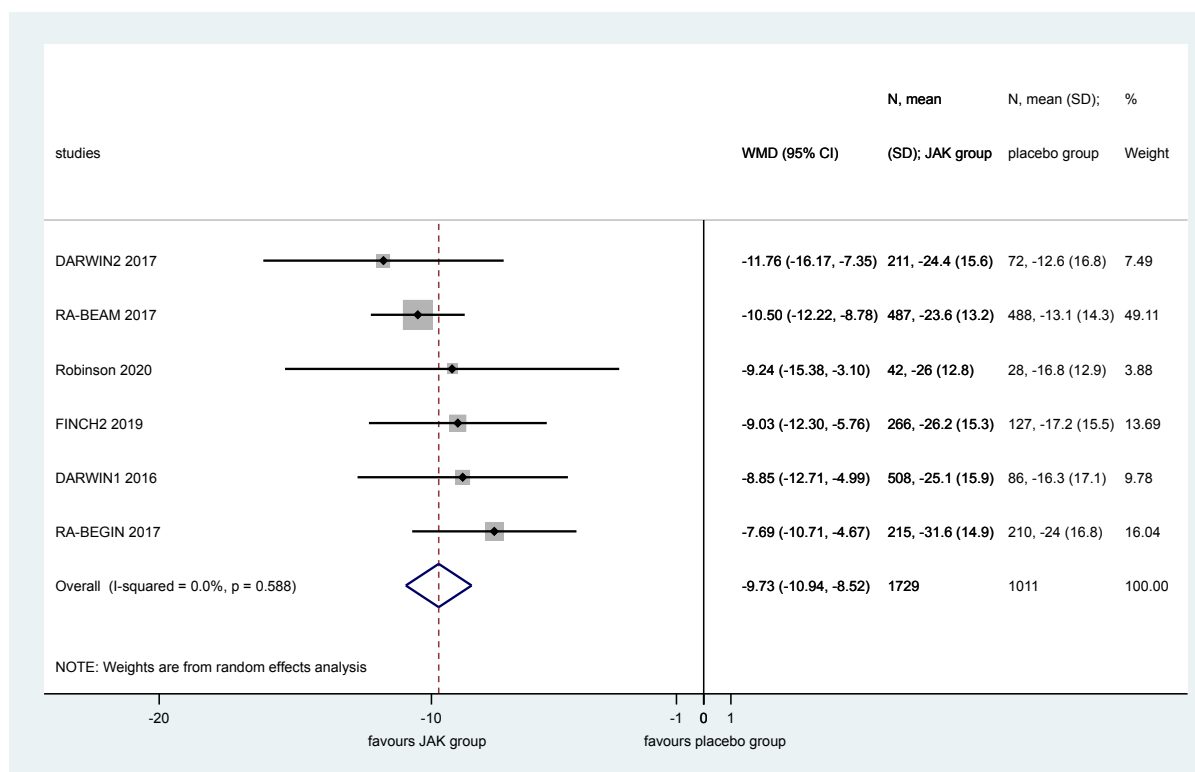
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
Tanaka 2011	-25.41 (-36.27, -14.55)	95, -41.3 (21.1)	24, -15.9 (25)	4.47
RAJ4 2019	-20.87 (-24.80, -16.94)	334, -33 (21)	159, -12.1 (20.8)	6.23
Kremer 2009	-20.38 (-28.73, -12.03)	166, -42.8 (23.6)	50, -22.4 (27.2)	5.16
SELECT-COMPARE 2019	-18.24 (-21.79, -14.69)	651, -45.5 (32.6)	651, -27.2 (32.7)	6.29
RAJ3 2019	-17.28 (-24.06, -10.50)	188, -25.3 (27.1)	88, -8.02 (26.6)	5.58
ORAL Solo 2012	-15.52 (-20.76, -10.28)	467, -34 (23.3)	109, -18.5 (25.5)	5.95
FINCH2 2019	-14.12 (-16.82, -11.42)	270, -41.5 (13.1)	128, -27.4 (12.7)	6.43
SELECT-BEYOND 2018	-13.95 (-18.54, -9.36)	329, -39.7 (24.5)	169, -25.8 (24.9)	6.10
RA-BEAM 2017	-13.80 (-16.67, -10.93)	487, -38.1 (22.2)	488, -24.3 (23.5)	6.40
Fleischmann 2012	-13.48 (-19.69, -7.27)	244, -33.1 (21.4)	46, -19.6 (19.4)	5.72
ORAL Scan 2013	-12.80 (-17.09, -8.51)	435, -36.6 (17)	61, -23.8 (15.9)	6.16
BALANCE II 2016	-11.48 (-17.18, -5.78)	246, -39.5 (18.5)	47, -28 (18.2)	5.85
RA-BEGIN 2017	-11.30 (-15.76, -6.84)	215, -49.7 (22)	210, -38.4 (24.8)	6.12
BALANCE I 2016	-9.74 (-16.01, -3.47)	217, -39.3 (21)	54, -29.6 (21.1)	5.70
Kremer 2012	-9.03 (-14.93, -3.13)	387, -24.1 (21.6)	60, -15.1 (21.7)	5.80
Kivitz 2016	-7.21 (-14.45, 0.03)	291, -32.4 (26.1)	67, -25.2 (27.5)	5.46
Robinson 2020	-1.95 (-2.91, -0.99)	42, -4.82 (1.19)	28, -2.87 (2.41)	6.59
Overall (I-squared = 94.8%, p = 0.000)	-13.69 (-17.72, -9.65)	5064	2439	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours placebo group

S83. Forest plot of studies comparing the LSM change from baseline in PGA measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and placebo within 6 months (PGA VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
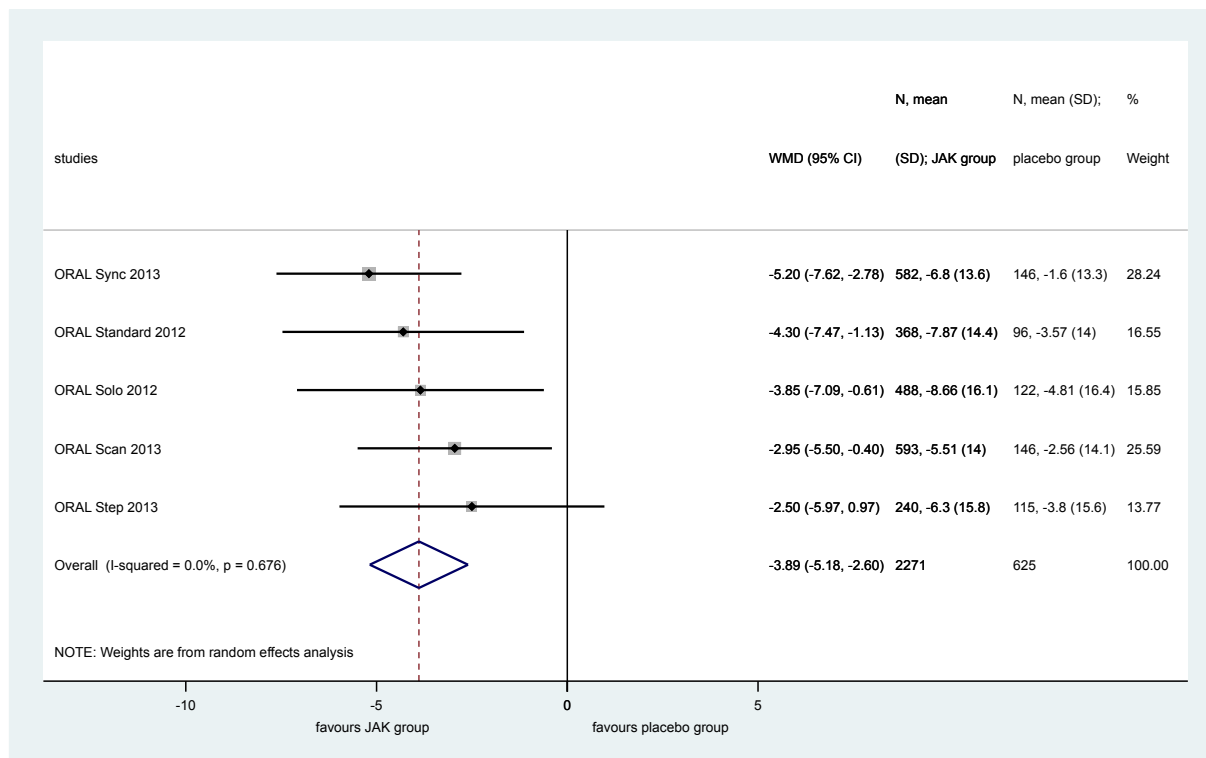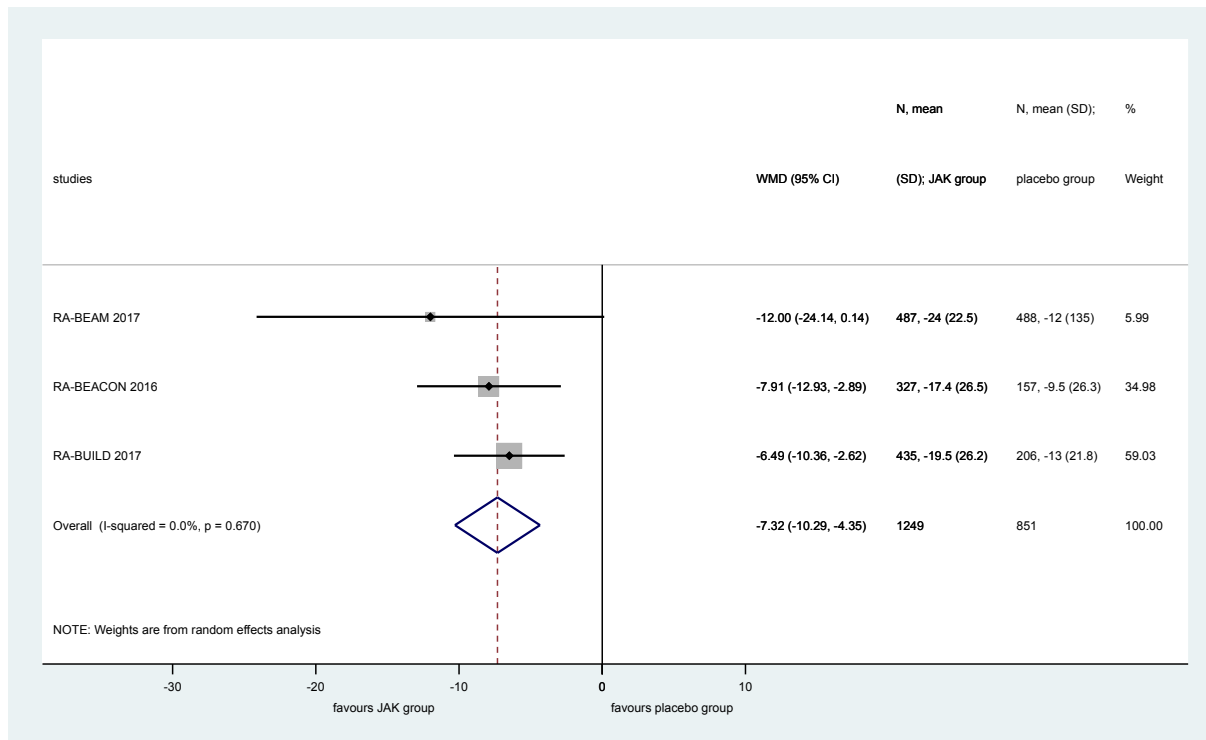
S84. Forest plot of studies comparing the LSM change from baseline in PtGA measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and placebo within 6 months (PtGA VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
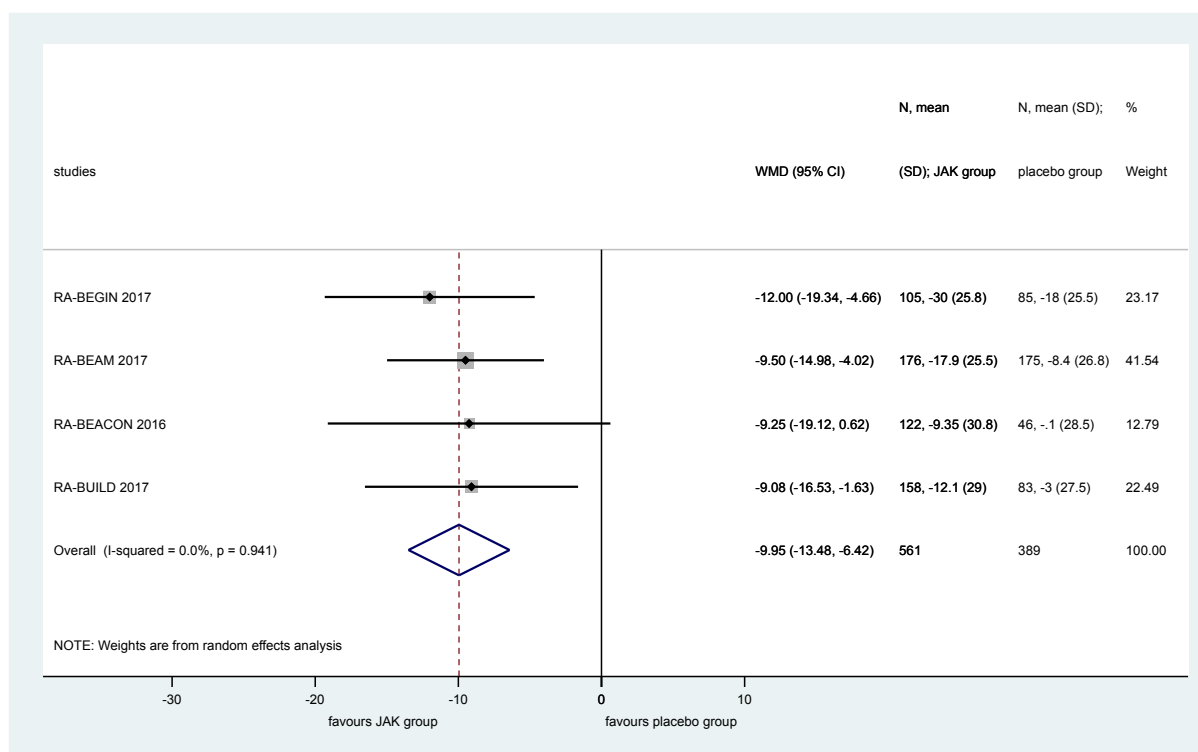
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
Tanaka 2011	-8.33 (-11.24, -5.42)	95, -10.3 (8.5)	24, -2 (5.88)	3.26
DARWIN2 2017	-6.29 (-8.97, -3.61)	211, -10.4 (8.93)	72, -4.1 (10.4)	3.64
ORAL Solo 2012	-5.00 (-6.92, -3.08)	466, -10.1 (9.61)	109, -5.12 (9.09)	5.31
RAJ4 2019	-4.79 (-5.94, -3.64)	334, -6.99 (6.11)	159, -2.2 (6.07)	7.68
SELECT-COMPARE 2019	-4.43 (-7.41, -1.45)	651, -11.6 (37.7)	651, -7.14 (9.23)	3.15
Fleischmann 2012	-4.22 (-6.70, -1.74)	247, -9.09 (8.27)	46, -4.87 (7.8)	4.01
Kremer 2009	-4.14 (-7.79, -0.49)	168, -13.3 (10)	50, -9.2 (12)	2.34
RA-BEAM 2017	-3.80 (-4.82, -2.78)	487, -9.5 (7.3)	488, -5.7 (8.9)	8.10
SELECT-BEYOND 2018	-3.79 (-5.52, -2.06)	329, -10.4 (9.86)	169, -6.6 (9.05)	5.82
RAJ3 2019	-3.63 (-5.31, -1.95)	188, -6.23 (6.64)	88, -2.6 (6.61)	5.99
ORAL-Scan 2013	-3.45 (-4.98, -1.92)	435, -9.96 (6.19)	62, -6.51 (5.67)	6.44
BALANCE I 2016	-3.08 (-5.54, -0.62)	217, -9.48 (8.26)	55, -6.4 (8.32)	4.05
Kremer 2012	-2.79 (-4.46, -1.12)	389, -9.14 (6.2)	61, -6.35 (6.17)	6.01
RA-BEGIN 2017	-2.70 (-4.57, -0.83)	215, -13.1 (9.3)	210, -10.4 (10.3)	5.44
DARWIN1 2016	-2.62 (-4.53, -0.71)	508, -10.2 (8.89)	86, -7.6 (8.25)	5.33
BALANCE II 2016	-2.56 (-4.47, -0.65)	246, -12.6 (6.08)	47, -10 (6.13)	5.32
FINCH2 2019	-2.45 (-3.42, -1.48)	276, -10.5 (4.51)	129, -8.02 (4.71)	8.27
Robinson 2020	-1.80 (-3.96, 0.36)	42, -8.43 (4.33)	28, -6.63 (4.65)	4.69
Kivitz 2016	-0.55 (-2.53, 1.43)	291, -7.55 (7.64)	67, -7 (7.4)	5.15
Overall (I-squared = 57.0%, p = 0.001)	-3.54 (-4.17, -2.90)	5795	2601	100.00

NOTE: Weights are from random effects analysis

favours JAK group | favours placebo group

S85. Forest plot of studies comparing the LSM change from baseline in the number of swollen joint counts between patients treated with JAK inhibitors and placebo within 6 months (Swollen Joint Count)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
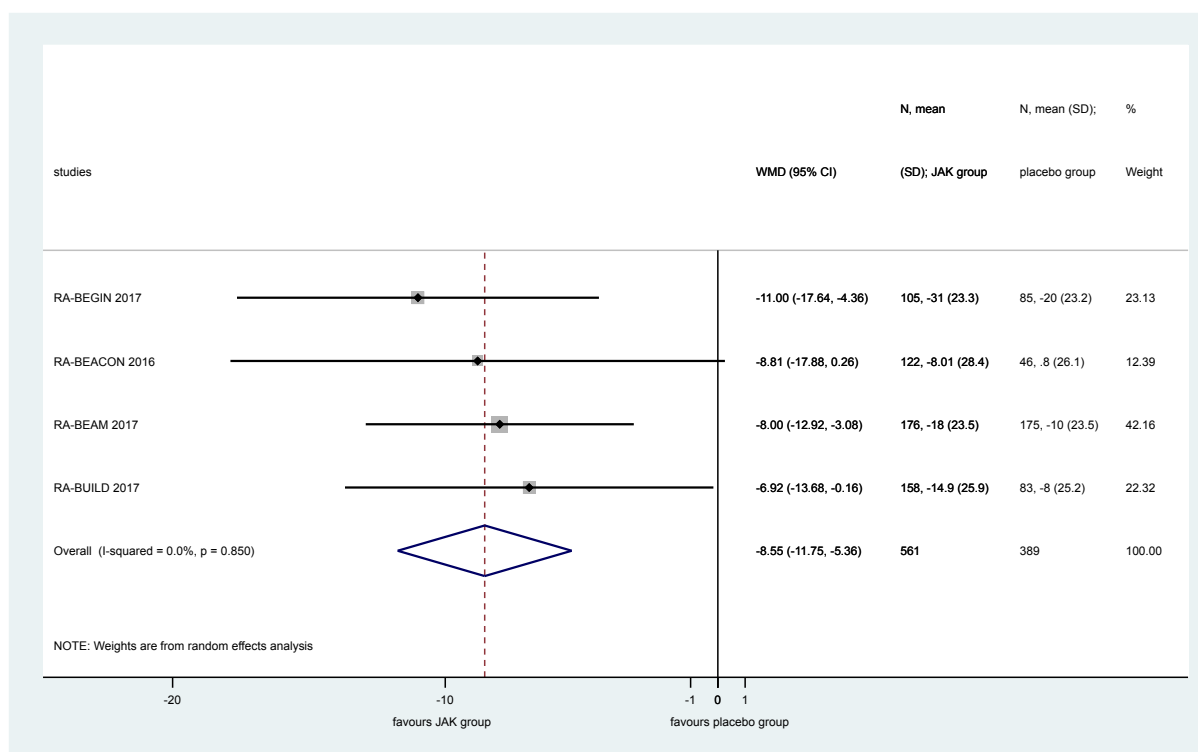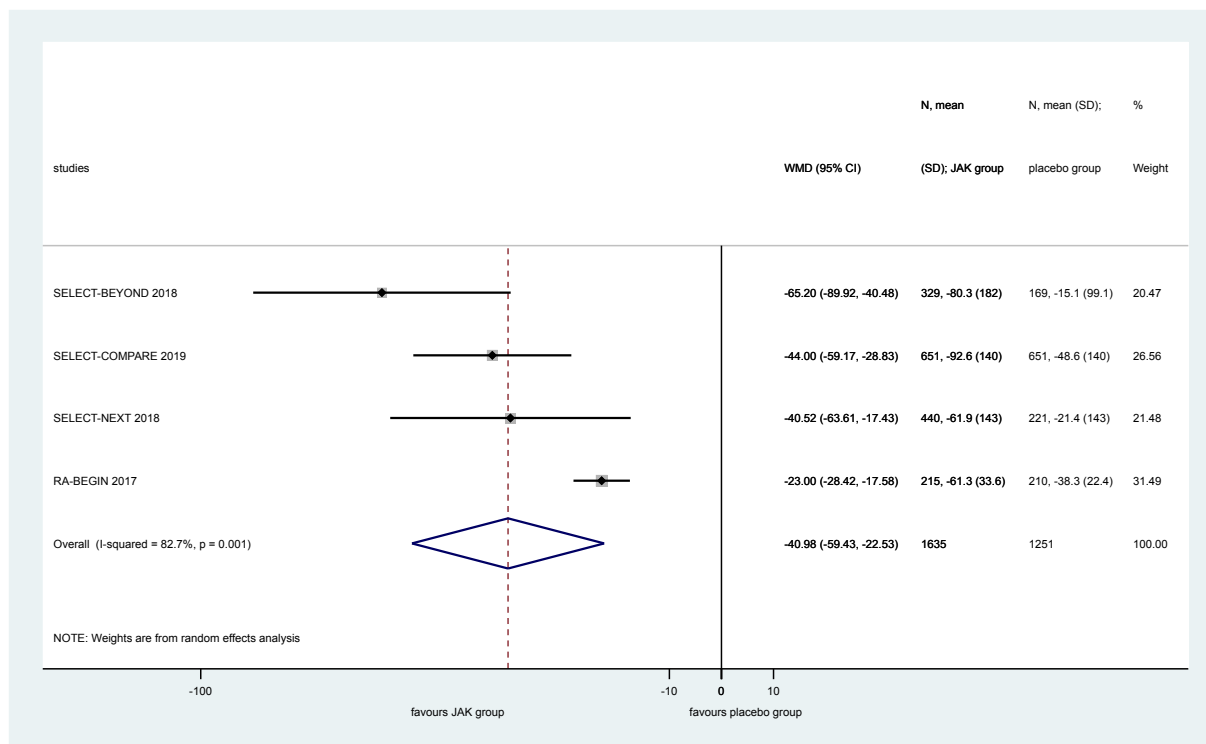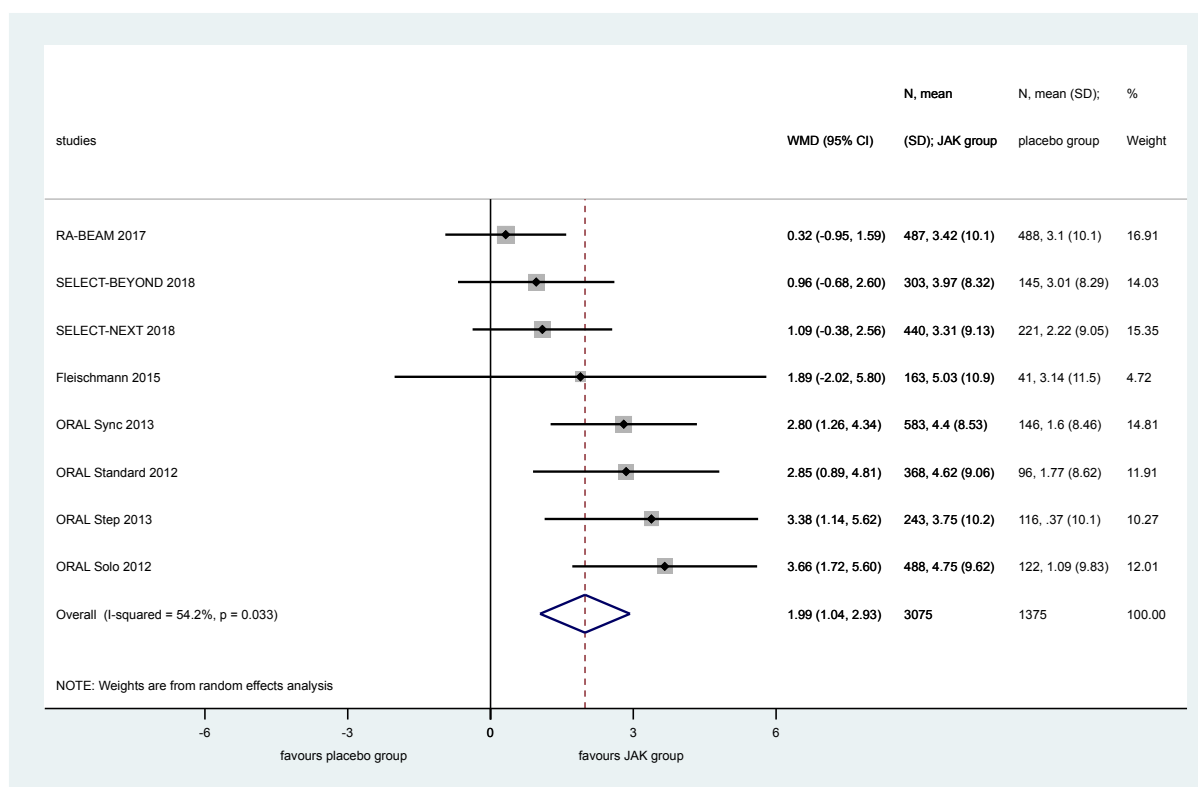
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
DARWIN2 2017	-9.23 (-12.58, -5.88)	211, -15 (12.4)	72, -5.8 (12.6)	3.17
Kremer 2009	-8.26 (-13.11, -3.41)	168, -20.2 (13.5)	50, -11.9 (15.9)	1.62
SELECT-BEYOND 2018	-7.36 (-9.85, -4.87)	329, -16.2 (13)	169, -8.82 (13.6)	5.15
ORAL Solo 2012	-6.73 (-9.80, -3.66)	466, -15.8 (13.9)	109, -9.06 (14.9)	3.67
SELECT-COMPARE 2019	-6.72 (-8.32, -5.12)	651, -16.4 (15)	651, -9.7 (14.4)	9.40
Tanaka 2011	-6.66 (-10.18, -3.14)	95, -11.8 (9.6)	24, -5.1 (7.35)	2.90
RAJ4 2019	-6.44 (-7.88, -5.00)	334, -8.34 (7.59)	159, -1.9 (7.66)	10.54
DARWIN1 2016	-6.00 (-8.86, -3.14)	508, -15.2 (12.5)	86, -9.2 (12.5)	4.12
BALANCE I 2016	-5.95 (-9.46, -2.44)	217, -15.3 (11.9)	55, -9.3 (11.8)	2.92
RA-BEAM 2017	-5.50 (-7.01, -3.99)	487, -13.2 (11.9)	488, -7.7 (12.2)	10.01
Kremer 2012	-4.92 (-7.41, -2.43)	389, -12.6 (9.16)	61, -7.63 (9.22)	5.15
FINCH2 2019	-4.88 (-6.23, -3.53)	276, -17 (6.51)	129, -12.1 (6.46)	11.26
RAJ3 2019	-4.69 (-6.77, -2.61)	188, -8.19 (8.17)	88, -3.5 (8.26)	6.68
Robinson 2020	-4.39 (-7.37, -1.41)	42, -10.3 (6.39)	28, -5.89 (6.12)	3.85
Fleischmann 2012	-4.12 (-7.50, -0.74)	247, -15.1 (11.3)	46, -11 (10.6)	3.10
RA-BEGIN 2017	-4.00 (-6.76, -1.24)	215, -20.2 (13.7)	210, -16.2 (15.3)	4.36
ORAL Scan 2013	-3.78 (-6.13, -1.43)	435, -14 (9.52)	62, -10.3 (8.74)	5.60
BALANCE II 2016	-3.70 (-7.15, -0.25)	246, -18.1 (11)	47, -14.4 (11.1)	3.00
Kivitz 2016	-2.33 (-5.49, 0.83)	291, -12.1 (13.1)	67, -9.8 (11.6)	3.49
Overall (I-squared = 25.3%, p = 0.152)	-5.52 (-6.16, -4.88)	5795	2601	100.00

NOTE: Weights are from random effects analysis

S86. Forest plot of studies comparing the LSM change from baseline in the number of tender joint counts between patients treated with JAK inhibitors and placebo within 6 months (Tender Joint Count)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
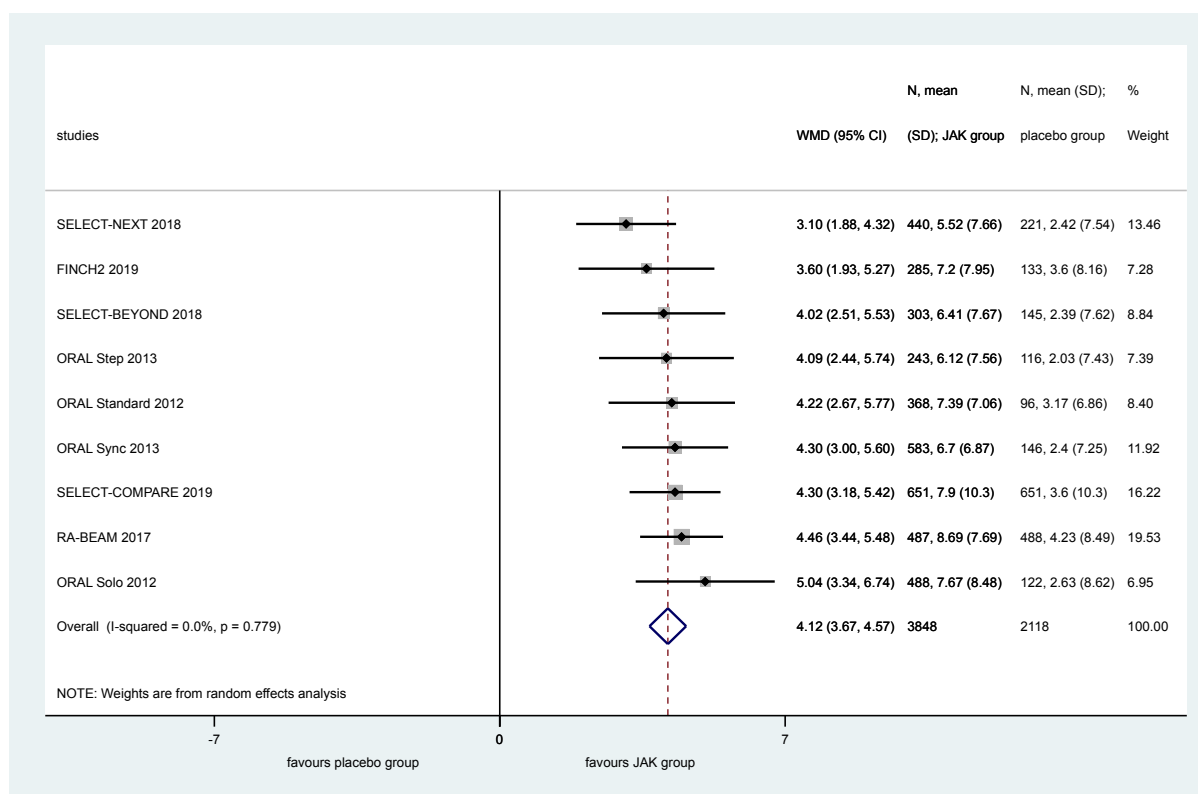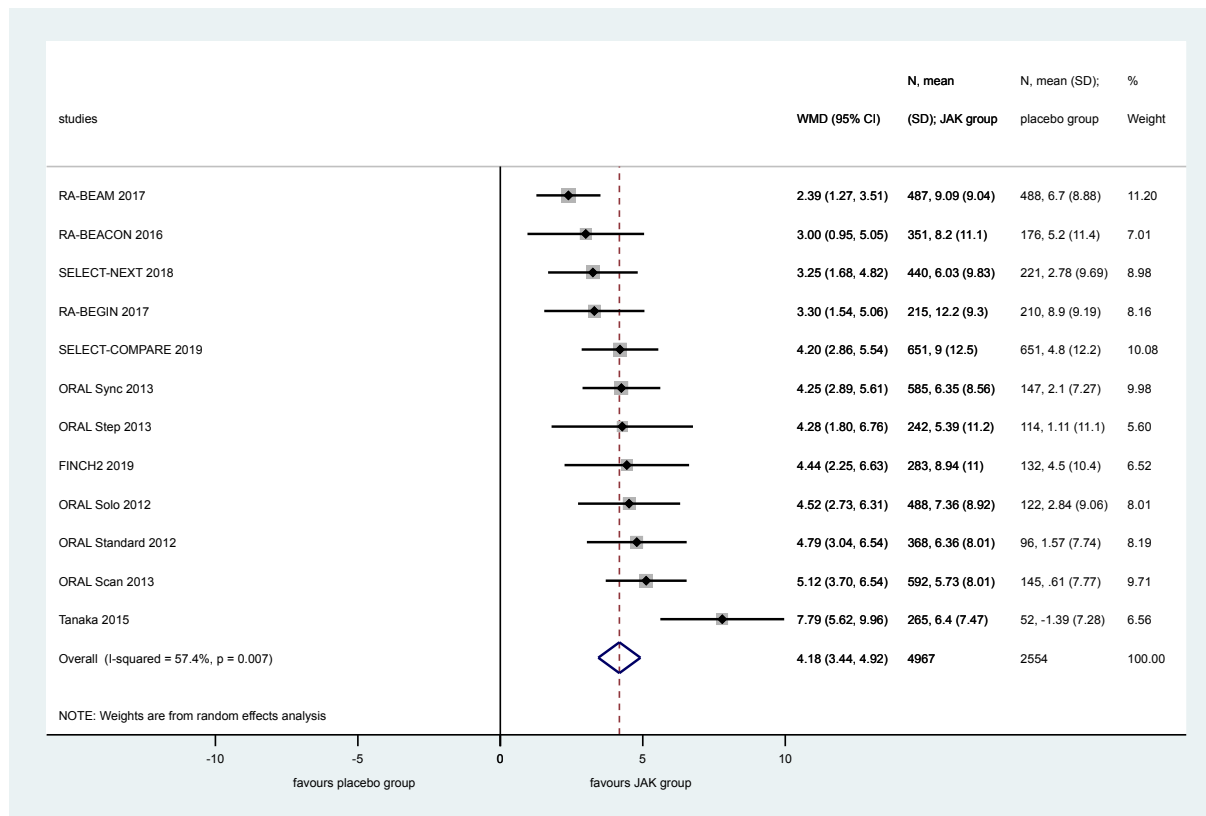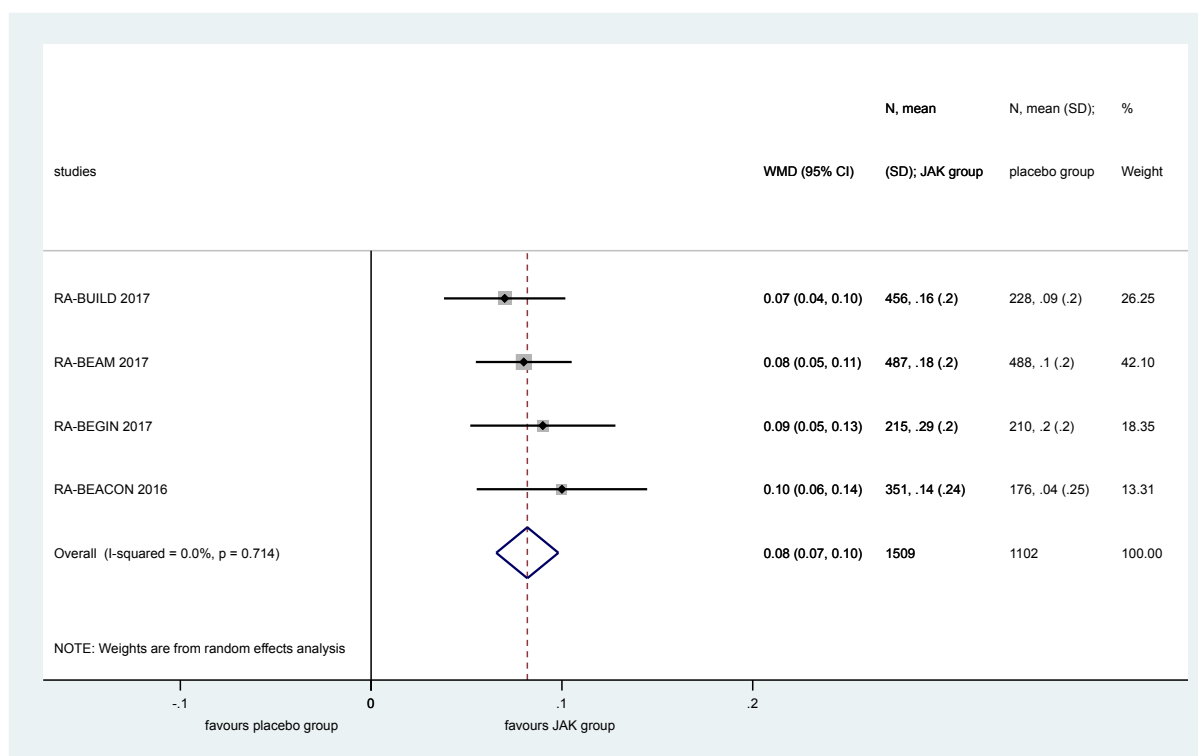
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
Fleischmann 2012	-27.11 (-32.62, -21.60)	240, -13.1 (18)	45, 14.1 (17.2)	3.86
Kremer 2009	-18.55 (-25.66, -11.44)	135, -17.9 (29.1)	41, .7 (16.8)	2.94
DARWIN1 2016	-17.29 (-22.19, -12.39)	508, -14.6 (26)	86, 2.67 (20.6)	4.28
RA-BEAM 2017	-12.57 (-15.42, -9.72)	487, -14.8 (23.6)	488, -2.25 (21.7)	5.88
SELECT-COMPARE 2019	-12.05 (-14.23, -9.87)	651, -13.5 (20.1)	651, -1.43 (20.1)	6.39
ORAL Scan 2013	-11.89 (-15.28, -8.50)	435, -11.1 (13.2)	63, .82 (12.8)	5.45
Tanaka 2011	-11.08 (-20.42, -1.74)	95, -14.8 (23.7)	24, -3.7 (20.1)	2.05
Kremer 2012	-10.71 (-14.26, -7.16)	368, -7.67 (13)	60, 3.04 (13)	5.32
ORAL Solo 2012	-10.48 (-13.21, -7.75)	465, -13.8 (13.1)	108, -3.3 (13)	5.98
SELECT-BEYOND 2018	-9.96 (-12.55, -7.37)	329, -11.2 (13.7)	169, -1.28 (14.1)	6.09
FINCH2 2019	-9.84 (-11.58, -8.10)	275, -11.6 (8.27)	129, -1.8 (8.32)	6.69
BALANCE II 2016	-8.42 (-11.66, -5.18)	246, -8.82 (10.3)	47, -.4 (10.4)	5.57
BALANCE I 2016	-7.41 (-12.06, -2.76)	217, -7.81 (15.7)	55, -.4 (15.7)	4.46
RA-BEGIN 2017	-6.67 (-11.34, -2.00)	215, -18.4 (27.9)	210, -11.7 (20.8)	4.44
DARWIN2 2017	-1.72 (-11.31, 7.87)	211, -10.4 (34.6)	72, -8.71 (36.3)	1.97
RAJ4 2019	-1.55 (-1.86, -1.24)	334, -1.52 (1.62)	158, .03 (1.62)	7.26
RAJ3 2019	-1.27 (-1.95, -0.59)	188, -.91 (2.74)	89, .36 (2.69)	7.19
Kivitz 2016	-0.51 (-0.79, -0.23)	290, -.36 (1.19)	67, .15 (1.01)	7.27
Robinson 2020	-0.27 (-1.65, 1.11)	42, -.55 (4.17)	28, -.28 (1.48)	6.90
Overall (I-squared = 97.1%, p = 0.000)	-8.57 (-10.14, -6.99)	5731	2590	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours placebo group

S87. Forest plot of studies comparing the LSM change from baseline in CRP value (mg/l) between patients treated with JAK inhibitors and placebo within 6 months (CRP)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
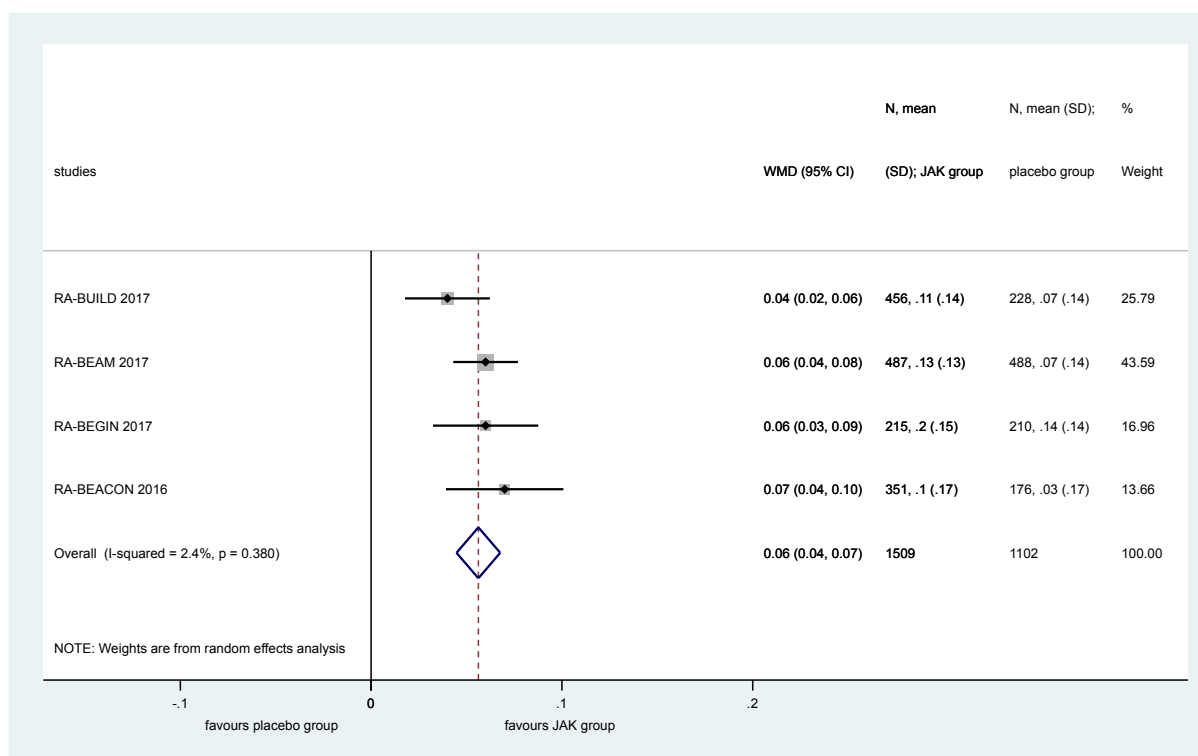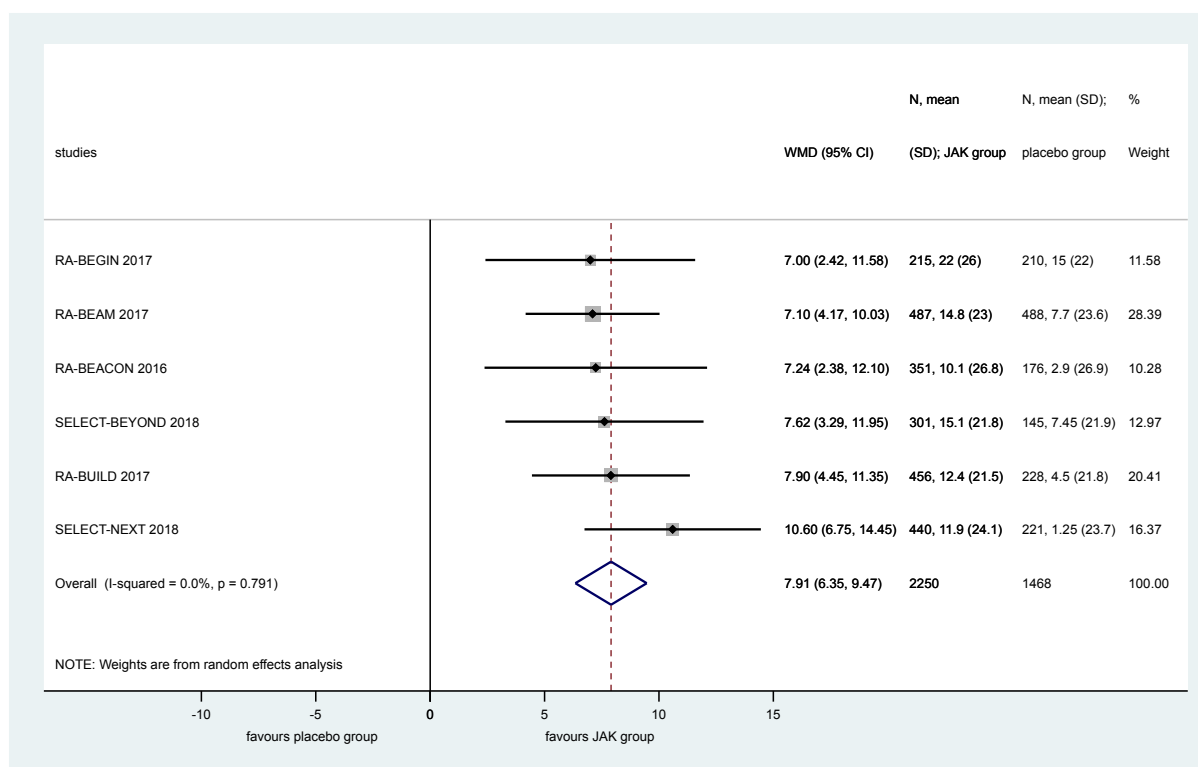
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
Tanaka 2015	-28.08 (-32.83, -23.33)	270, -20.1 (19.2)	52, 7.93 (15.3)	14.16
RAJ4 2019	-18.93 (-22.27, -15.59)	334, -21.4 (17.8)	159, -2.44 (17.7)	15.08
RAJ3 2019	-15.30 (-21.92, -8.68)	188, -16.2 (26.8)	88, -.89 (25.9)	12.69
RA-BEAM 2017	-12.40 (-15.32, -9.48)	487, -17.9 (25.5)	488, -5.5 (20.8)	15.32
Kivitz 2016	-9.38 (-13.52, -5.24)	288, -14.1 (17.1)	67, -4.7 (15.2)	14.58
ORAL Solo 2012	-8.80 (-13.18, -4.42)	442, -16.5 (20.6)	103, -7.7 (20.4)	14.42
RA-BEGIN 2017	-6.30 (-11.58, -1.02)	215, -23.4 (27.3)	210, -17.1 (28.2)	13.76
Overall (I-squared = 90.3%, p = 0.000)	-14.17 (-19.28, -9.07)	2224	1167	100.00

NOTE: Weights are from random effects analysis

favours JAK group — favours placebo group

S88. Forest plot of studies comparing the LSM change from baseline in ESR value (mm/hour) between patients treated with JAK inhibitors and placebo within 6 months (ESR)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
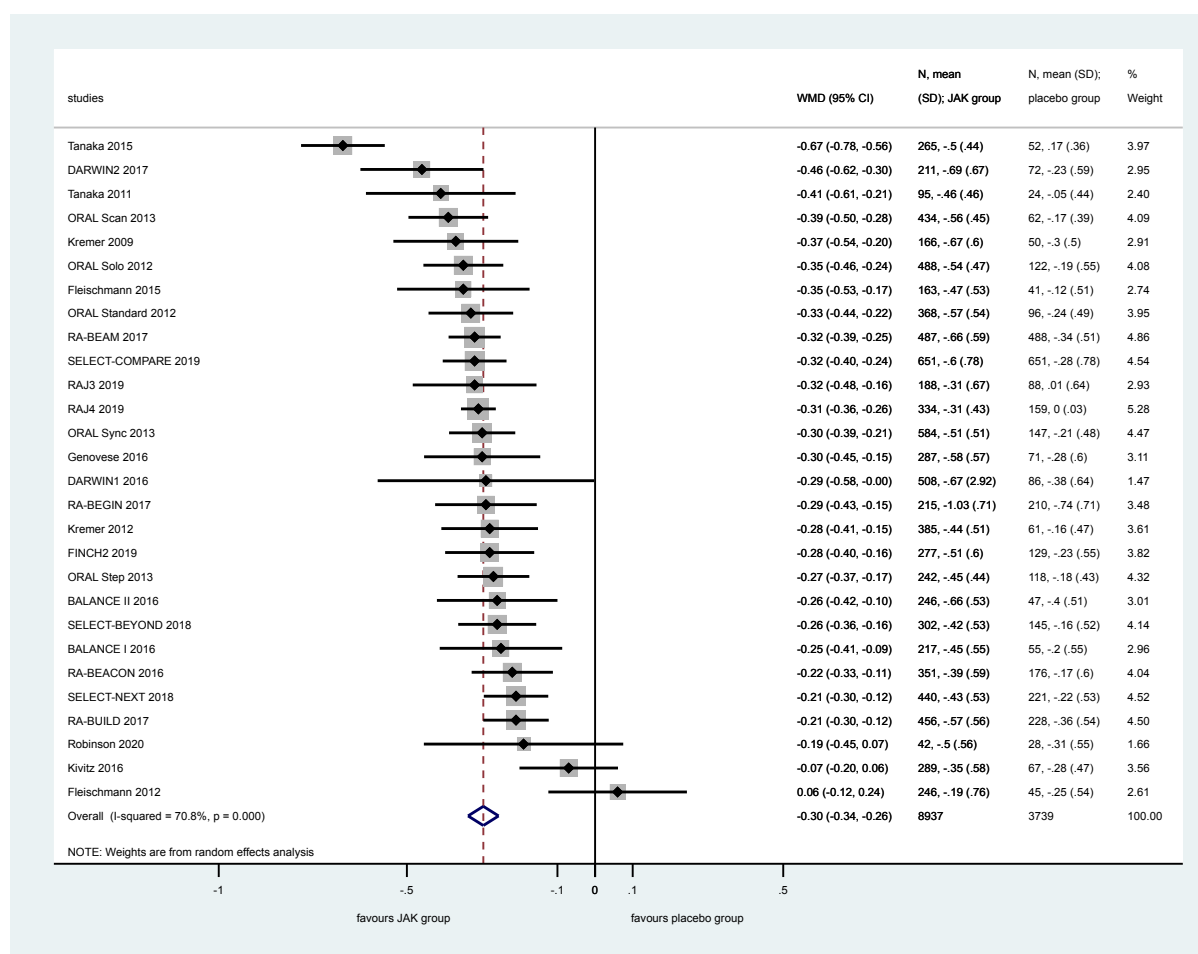
S89. Forest plot of studies comparing the LSM change from baseline in the score on HAQ-DI between patients treated with JAK inhibitors and placebo within 6 months (HAQ-DI difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
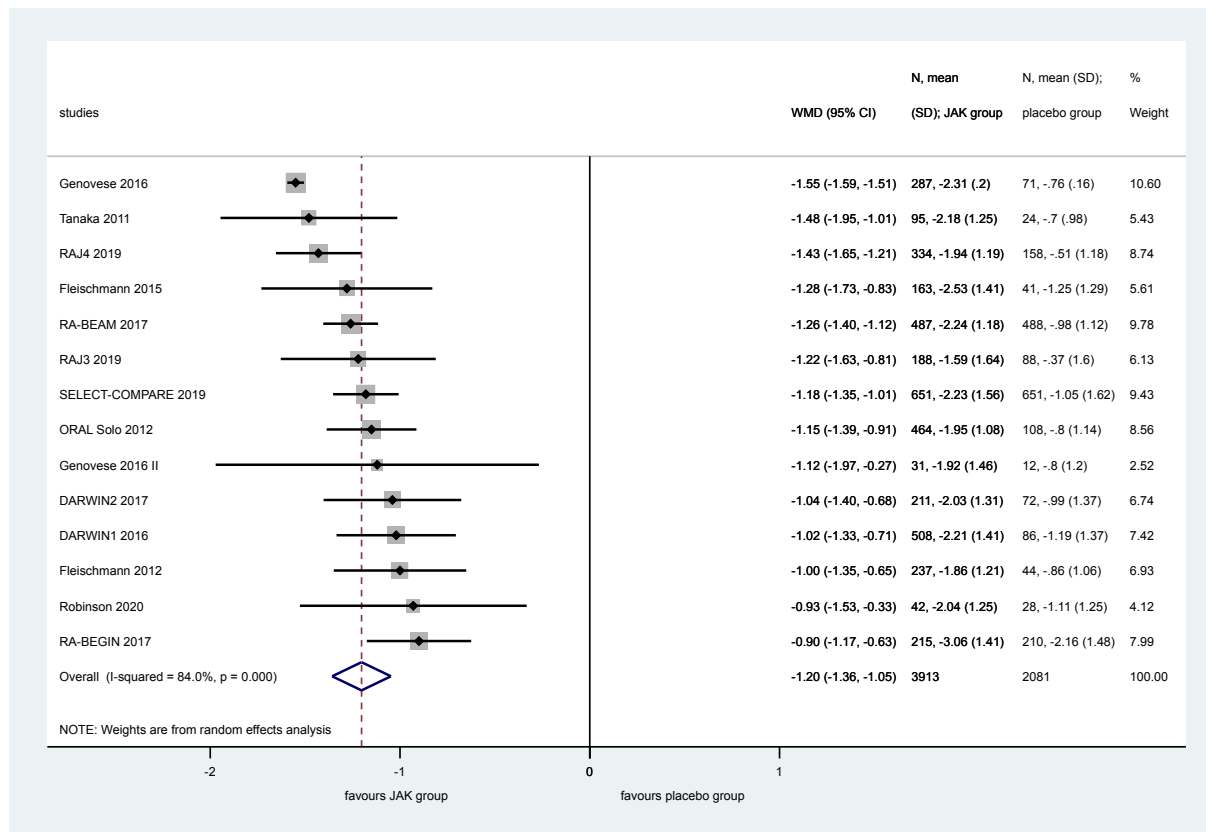
S90. Forest plot of studies comparing the LSM change from baseline in the score on DAS28-CRP between patients treated with JAK inhibitors and placebo within 6 months (DAS28-CRP difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
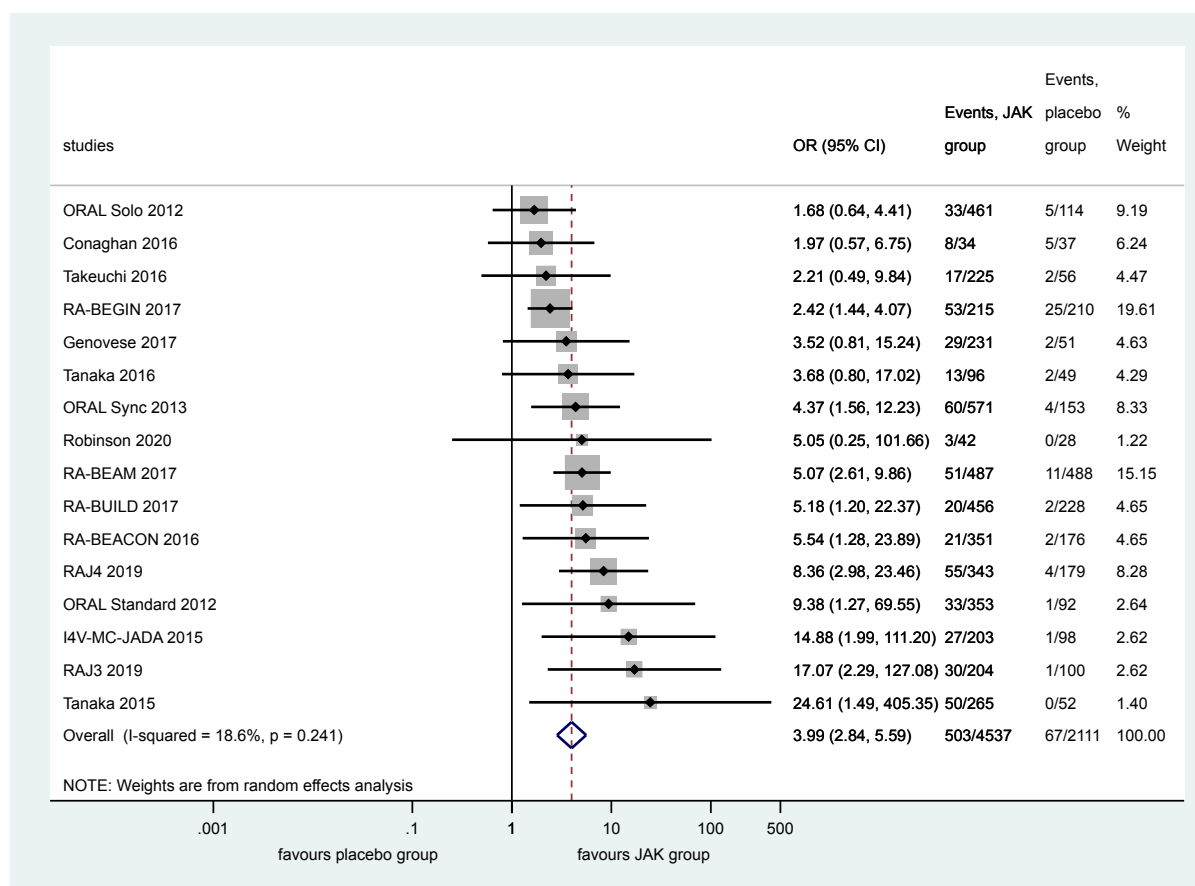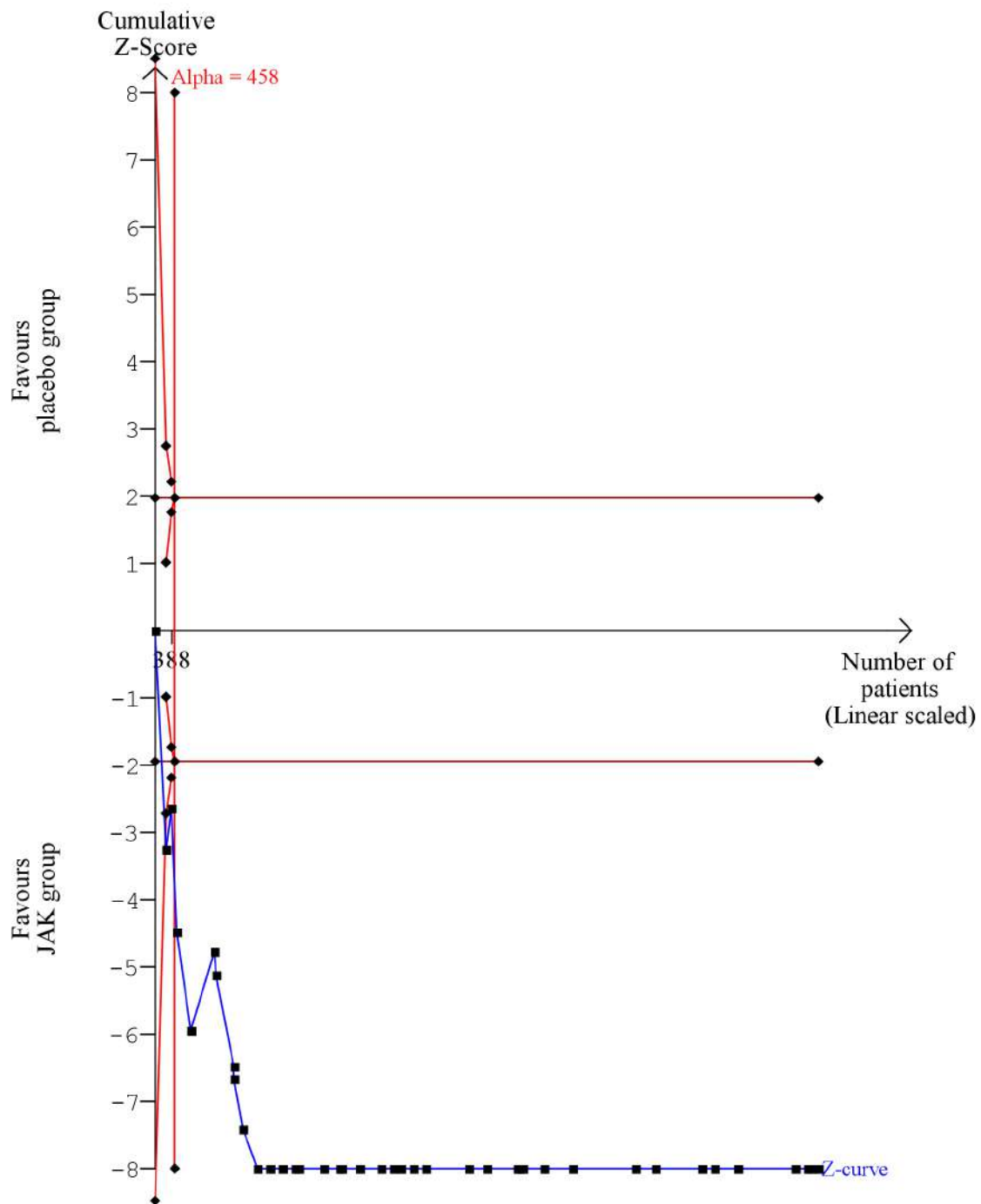
S91. Forest plot of studies comparing the LSM change from baseline in the score on DAS28-ESR between patients treated with JAK inhibitors and placebo within 6 months (DAS28-ESR difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
Genovese 2016	-13.73 (-17.18, -10.28)	287, -24.9 (12.5)	71, -11.2 (13.5)	15.26
DARWIN2 2017	-11.64 (-15.86, -7.42)	211, -23.3 (15.2)	72, -11.7 (15.9)	12.32
RA-BEAM 2017	-9.18 (-10.85, -7.51)	487, -22 (12.7)	488, -12.9 (13.9)	24.31
FINCH2 2019	-7.68 (-10.85, -4.51)	269, -25 (14.7)	128, -17.3 (15.2)	16.53
DARWIN1 2016	-7.07 (-10.92, -3.22)	508, -23.7 (15.4)	86, -16.6 (17.1)	13.67
RA-BEGIN 2017	-6.97 (-9.85, -4.09)	215, -29.8 (14.1)	210, -22.8 (16.1)	17.92
Overall (I-squared = 59.2%, p = 0.032)	-9.24 (-11.18, -7.31)	1977	1055	100.00

NOTE: Weights are from random effects analysis

S92. Forest plot of studies comparing the LSM change from baseline in the score on CDAI between patients treated with JAK inhibitors and placebo within 6 months (CDAI difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
DARWIN2 2017	-11.76 (-16.17, -7.35)	211, -24.4 (15.6)	72, -12.6 (16.8)	7.49
RA-BEAM 2017	-10.50 (-12.22, -8.78)	487, -23.6 (13.2)	488, -13.1 (14.3)	49.11
Robinson 2020	-9.24 (-15.38, -3.10)	42, -26 (12.8)	28, -16.8 (12.9)	3.88
FINCH2 2019	-9.03 (-12.30, -5.76)	266, -26.2 (15.3)	127, -17.2 (15.5)	13.69
DARWIN1 2016	-8.85 (-12.71, -4.99)	508, -25.1 (15.9)	86, -16.3 (17.1)	9.78
RA-BEGIN 2017	-7.69 (-10.71, -4.67)	215, -31.6 (14.9)	210, -24 (16.8)	16.04
Overall (I-squared = 0.0%, p = 0.588)	-9.73 (-10.94, -8.52)	1729	1011	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours placebo group

S93. Forest plot of studies comparing the LSM change from baseline in the score on SDAI between patients treated with JAK inhibitors and placebo within 6 months (SDAI difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
ORAL Sync 2013	-5.20 (-7.62, -2.78)	582, -6.8 (13.6)	146, -1.6 (13.3)	28.24
ORAL Standard 2012	-4.30 (-7.47, -1.13)	368, -7.87 (14.4)	96, -3.57 (14)	16.55
ORAL Solo 2012	-3.85 (-7.09, -0.61)	488, -8.66 (16.1)	122, -4.81 (16.4)	15.85
ORAL Scan 2013	-2.95 (-5.50, -0.40)	593, -5.51 (14)	146, -2.56 (14.1)	25.59
ORAL Step 2013	-2.50 (-5.97, 0.97)	240, -6.3 (15.8)	115, -3.8 (15.6)	13.77
Overall (I-squared = 0.0%, p = 0.676)	-3.89 (-5.18, -2.60)	2271	625	100.00

NOTE: Weights are from random effects analysis

S94. Forest plot of studies comparing the LSM change from baseline in the score on MOS-Sleep between patients treated with JAK inhibitors and placebo within 6 months (MOS-Sleep)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S95. Forest plot of studies comparing the LSM change from baseline in the score on WPAI assessing activity impairment between patients treated with JAK inhibitors and placebo within 6 months (WPAI AI)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
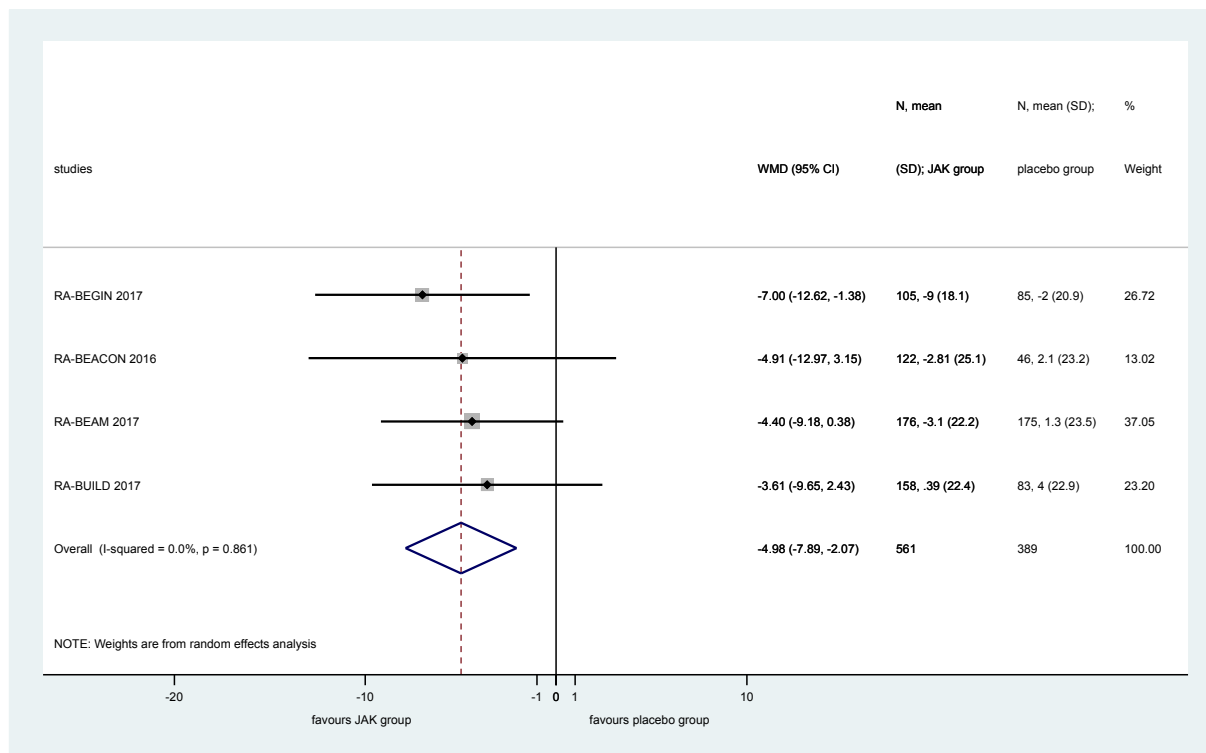
S96. Forest plot of studies comparing the LSM change from baseline in the score on WPAI assessing absenteeism (missed worked time) between patients treated with JAK inhibitors and placebo within 6 months (WPAI A)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S97. Forest plot of studies comparing the LSM change from baseline in the score on WPAI assessing overall work impairment (productivity loss) between patients treated with JAK inhibitors and placebo within 6 months (WPAI OWI)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BEGIN 2017	-11.00 (-17.64, -4.36)	105, -31 (23.3)	85, -20 (23.2)	23.13
RA-BEACON 2016	-8.81 (-17.88, 0.26)	122, -8.01 (28.4)	46, .8 (26.1)	12.39
RA-BEAM 2017	-8.00 (-12.92, -3.08)	176, -18 (23.5)	175, -10 (23.5)	42.16
RA-BUILD 2017	-6.92 (-13.68, -0.16)	158, -14.9 (25.9)	83, -8 (25.2)	22.32
Overall (I-squared = 0.0%, p = 0.850)	-8.55 (-11.75, -5.36)	561	389	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours placebo group

S98. Forest plot of studies comparing the LSM change from baseline in the score on WPAI assessing presenteeism (impairment while working) between patients treated with JAK inhibitors and placebo within 6 months (WPAI P)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
SELECT-BEYOND 2018	-65.20 (-89.92, -40.48)	329, -80.3 (182)	169, -15.1 (99.1)	20.47
SELECT-COMPARE 2019	-44.00 (-59.17, -28.83)	651, -92.6 (140)	651, -48.6 (140)	26.56
SELECT-NEXT 2018	-40.52 (-63.61, -17.43)	440, -61.9 (143)	221, -21.4 (143)	21.48
RA-BEGIN 2017	-23.00 (-28.42, -17.58)	215, -61.3 (33.6)	210, -38.3 (22.4)	31.49
Overall (I-squared = 82.7%, p = 0.001)	-40.98 (-59.43, -22.53)	1635	1251	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours placebo group

S99. Forest plot of studies comparing the LSM change from baseline in the duration of morning joint stiffness between patients treated with JAK inhibitors and placebo within 6 months (MJS duration)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BEAM 2017	0.32 (-0.95, 1.59)	487, 3.42 (10.1)	488, 3.1 (10.1)	16.91
SELECT-BEYOND 2018	0.96 (-0.68, 2.60)	303, 3.97 (8.32)	145, 3.01 (8.29)	14.03
SELECT-NEXT 2018	1.09 (-0.38, 2.56)	440, 3.31 (9.13)	221, 2.22 (9.05)	15.35
Fleischmann 2015	1.89 (-2.02, 5.80)	163, 5.03 (10.9)	41, 3.14 (11.5)	4.72
ORAL Sync 2013	2.80 (1.26, 4.34)	583, 4.4 (8.53)	146, 1.6 (8.46)	14.81
ORAL Standard 2012	2.85 (0.89, 4.81)	368, 4.62 (9.06)	96, 1.77 (8.62)	11.91
ORAL Step 2013	3.38 (1.14, 5.62)	243, 3.75 (10.2)	116, .37 (10.1)	10.27
ORAL Solo 2012	3.66 (1.72, 5.60)	488, 4.75 (9.62)	122, 1.09 (9.83)	12.01
Overall (I-squared = 54.2%, p = 0.033)	1.99 (1.04, 2.93)	3075	1375	100.00

NOTE: Weights are from random effects analysis

S100. Forest plot of studies comparing the LSM change from baseline in the score on SF-36 assessing the mental component score between patients treated with JAK inhibitors and placebo within 6 months (SF-36 MCS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
SELECT-NEXT 2018	3.10 (1.88, 4.32)	440, 5.52 (7.66)	221, 2.42 (7.54)	13.46
FINCH2 2019	3.60 (1.93, 5.27)	285, 7.2 (7.95)	133, 3.6 (8.16)	7.28
SELECT-BEYOND 2018	4.02 (2.51, 5.53)	303, 6.41 (7.67)	145, 2.39 (7.62)	8.84
ORAL Step 2013	4.09 (2.44, 5.74)	243, 6.12 (7.56)	116, 2.03 (7.43)	7.39
ORAL Standard 2012	4.22 (2.67, 5.77)	368, 7.39 (7.06)	96, 3.17 (6.86)	8.40
ORAL Sync 2013	4.30 (3.00, 5.60)	583, 6.7 (6.87)	146, 2.4 (7.25)	11.92
SELECT-COMPARE 2019	4.30 (3.18, 5.42)	651, 7.9 (10.3)	651, 3.6 (10.3)	16.22
RA-BEAM 2017	4.46 (3.44, 5.48)	487, 8.69 (7.69)	488, 4.23 (8.49)	19.53
ORAL Solo 2012	5.04 (3.34, 6.74)	488, 7.67 (8.48)	122, 2.63 (8.62)	6.95
Overall (I-squared = 0.0%, p = 0.779)	4.12 (3.67, 4.57)	3848	2118	100.00

NOTE: Weights are from random effects analysis

S101. Forest plot of studies comparing the LSM change from baseline in the score on SF-36 assessing the physical component score between patients treated with JAK inhibitors and placebo within 6 months (SF-36 PCS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BEAM 2017	2.39 (1.27, 3.51)	487, 9.09 (9.04)	488, 6.7 (8.88)	11.20
RA-BEACON 2016	3.00 (0.95, 5.05)	351, 8.2 (11.1)	176, 5.2 (11.4)	7.01
SELECT-NEXT 2018	3.25 (1.68, 4.82)	440, 6.03 (9.83)	221, 2.78 (9.69)	8.98
RA-BEGIN 2017	3.30 (1.54, 5.06)	215, 12.2 (9.3)	210, 8.9 (9.19)	8.16
SELECT-COMPARE 2019	4.20 (2.86, 5.54)	651, 9 (12.5)	651, 4.8 (12.2)	10.08
ORAL Sync 2013	4.25 (2.89, 5.61)	585, 6.35 (8.56)	147, 2.1 (7.27)	9.98
ORAL Step 2013	4.28 (1.80, 6.76)	242, 5.39 (11.2)	114, 1.11 (11.1)	5.60
FINCH2 2019	4.44 (2.25, 6.63)	283, 8.94 (11)	132, 4.5 (10.4)	6.52
ORAL Solo 2012	4.52 (2.73, 6.31)	488, 7.36 (8.92)	122, 2.84 (9.06)	8.01
ORAL Standard 2012	4.79 (3.04, 6.54)	368, 6.36 (8.01)	96, 1.57 (7.74)	8.19
ORAL Scan 2013	5.12 (3.70, 6.54)	592, 5.73 (8.01)	145, .61 (7.77)	9.71
Tanaka 2015	7.79 (5.62, 9.96)	265, 6.4 (7.47)	52, -1.39 (7.28)	6.56
Overall (I-squared = 57.4%, p = 0.007)	4.18 (3.44, 4.92)	4967	2554	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S102. Forest plot of studies comparing the LSM change from baseline in the score on FACIT-F between patients treated with JAK inhibitors and placebo within 6 months (FACIT-F)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BUILD 2017	0.07 (0.04, 0.10)	456, .16 (.2)	228, .09 (.2)	26.25
RA-BEAM 2017	0.08 (0.05, 0.11)	487, .18 (.2)	488, .1 (.2)	42.10
RA-BEGIN 2017	0.09 (0.05, 0.13)	215, .29 (.2)	210, .2 (.2)	18.35
RA-BEACON 2016	0.10 (0.06, 0.14)	351, .14 (.24)	176, .04 (.25)	13.31
Overall (I-squared = 0.0%, p = 0.714)	0.08 (0.07, 0.10)	1509	1102	100.00

NOTE: Weights are from random effects analysis

-.1   favours placebo group   0   .1 favours JAK group   .2

S103. Forest plot of studies comparing the LSM change from baseline in the score on EQ-5D assessed by the UK scoring algorithm between patients treated with JAK inhibitors and placebo within 6 months (EQ-5D UK)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BUILD 2017	0.04 (0.02, 0.06)	456, .11 (.14)	228, .07 (.14)	25.79
RA-BEAM 2017	0.06 (0.04, 0.08)	487, .13 (.13)	488, .07 (.14)	43.59
RA-BEGIN 2017	0.06 (0.03, 0.09)	215, .2 (.15)	210, .14 (.14)	16.96
RA-BEACON 2016	0.07 (0.04, 0.10)	351, .1 (.17)	176, .03 (.17)	13.66
Overall (I-squared = 2.4%, p = 0.380)	0.06 (0.04, 0.07)	1509	1102	100.00

NOTE: Weights are from random effects analysis

favours placebo group   favours JAK group

S104. Forest plot of studies comparing the LSM change from baseline in the score on EQ-5D assessed by the US scoring algorithm between patients treated with JAK inhibitors and placebo within 6 months (EQ-5D US)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); placebo group	% Weight
RA-BEGIN 2017	7.00 (2.42, 11.58)	215, 22 (26)	210, 15 (22)	11.58
RA-BEAM 2017	7.10 (4.17, 10.03)	487, 14.8 (23)	488, 7.7 (23.6)	28.39
RA-BEACON 2016	7.24 (2.38, 12.10)	351, 10.1 (26.8)	176, 2.9 (26.9)	10.28
SELECT-BEYOND 2018	7.62 (3.29, 11.95)	301, 15.1 (21.8)	145, 7.45 (21.9)	12.97
RA-BUILD 2017	7.90 (4.45, 11.35)	456, 12.4 (21.5)	228, 4.5 (21.8)	20.41
SELECT-NEXT 2018	10.60 (6.75, 14.45)	440, 11.9 (24.1)	221, 1.25 (23.7)	16.37
Overall (I-squared = 0.0%, p = 0.791)	7.91 (6.35, 9.47)	2250	1468	100.00

NOTE: Weights are from random effects analysis

S105. Forest plot of studies comparing the LSM change from baseline in the score on EQ-5D assessed by VAS between patients treated with JAK inhibitors and placebo within 6 months (EQ-5D VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
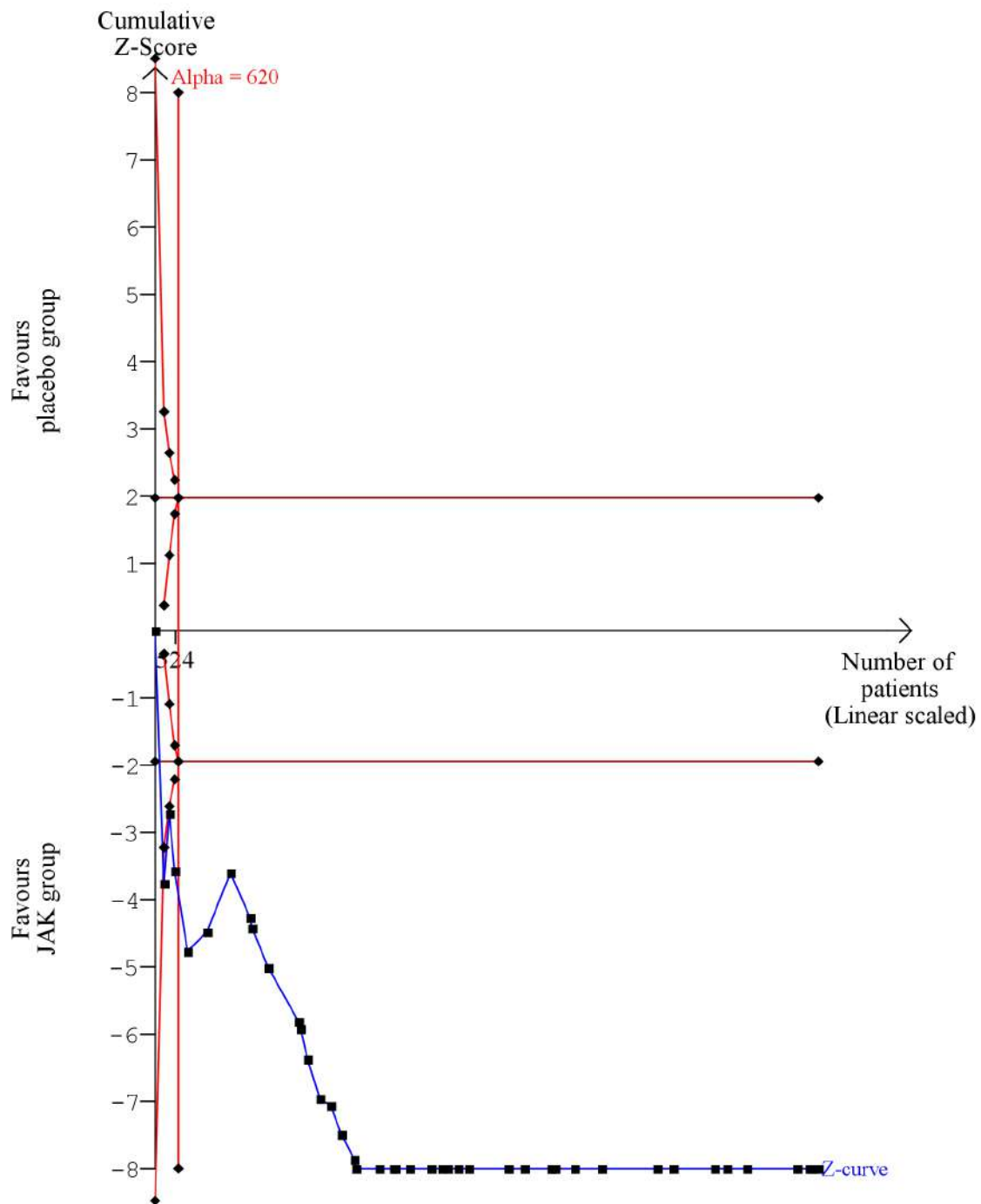
S106. TSA of ACR20 response (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.
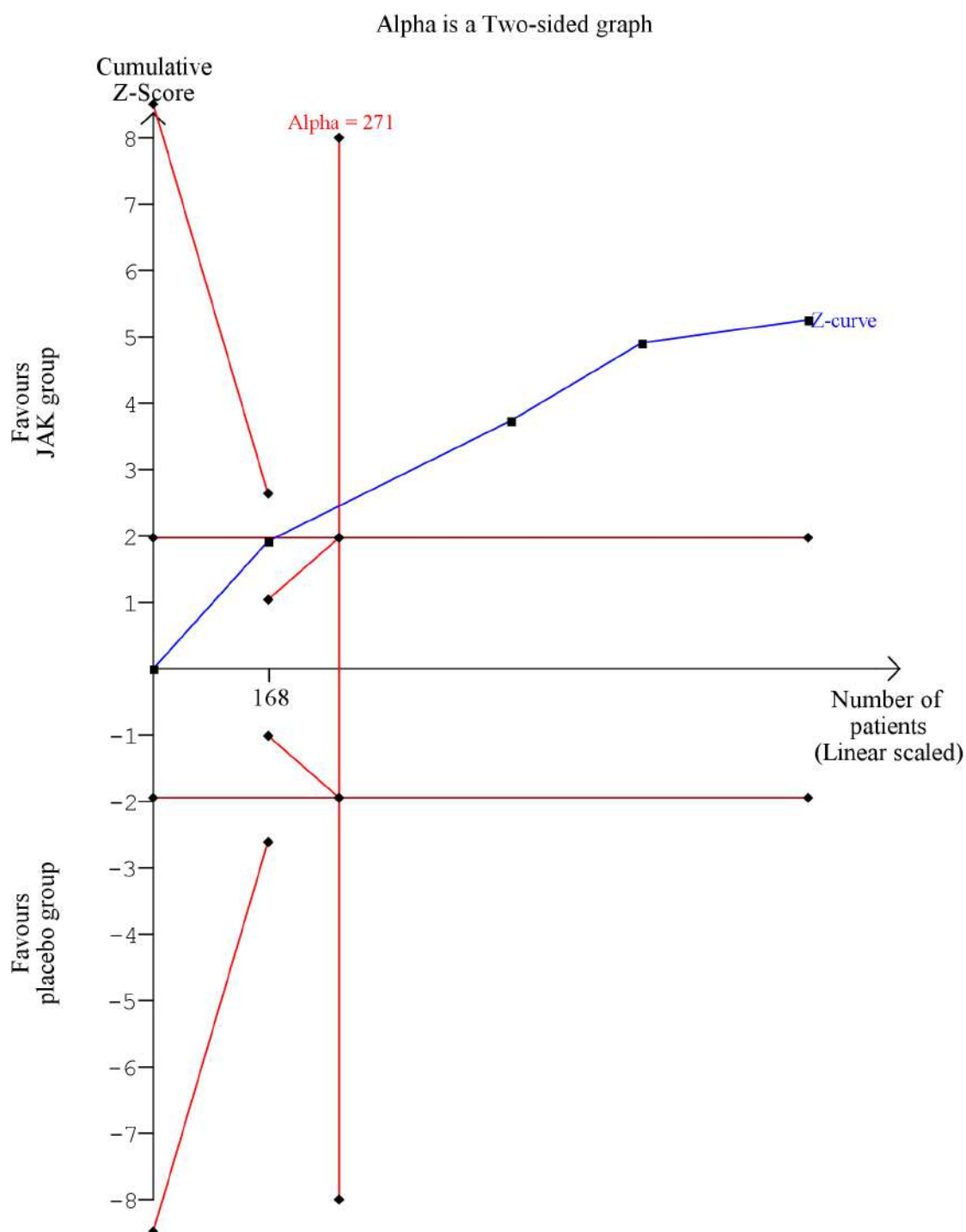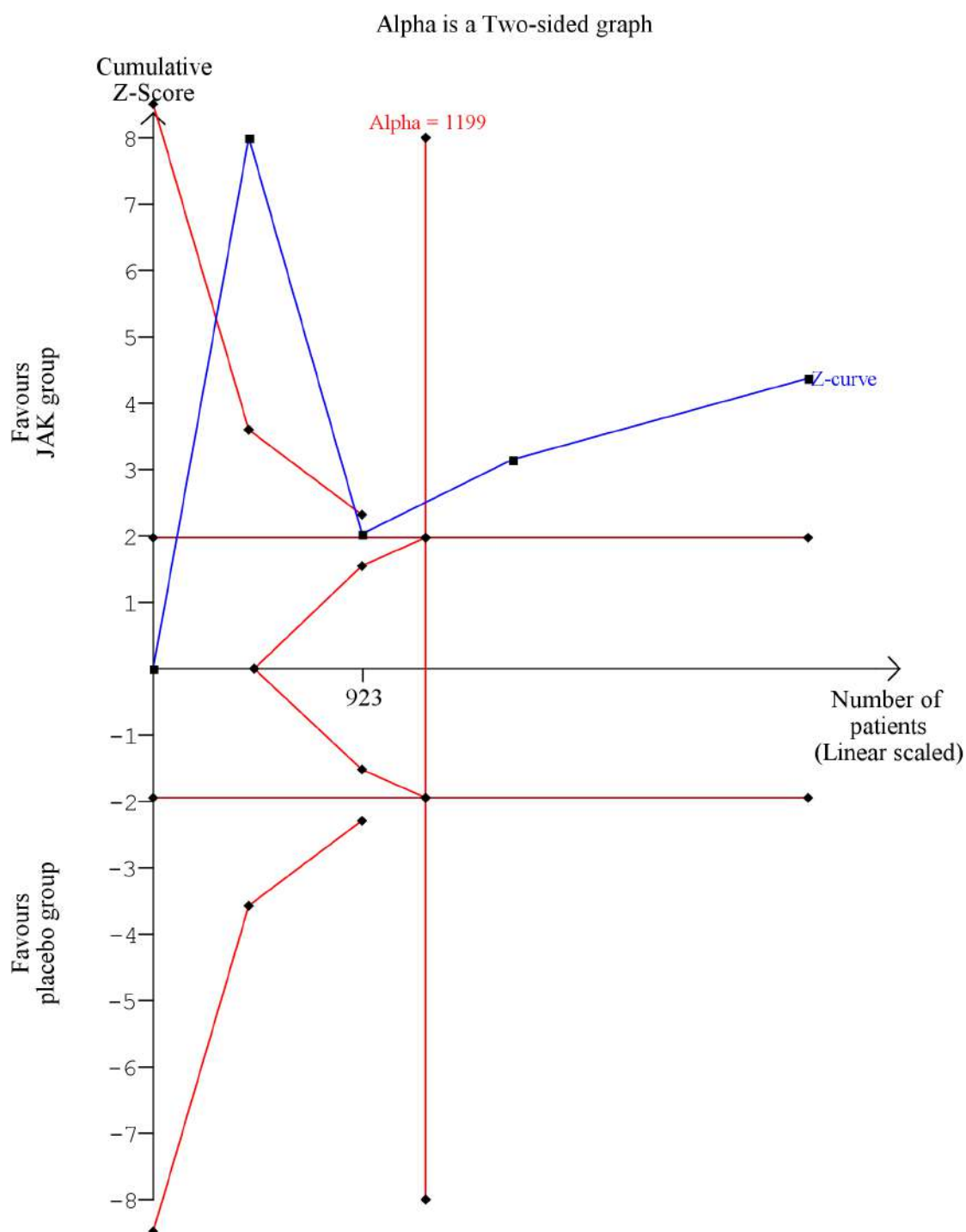
Alpha is a Two-sided graph

S107. TSA of ACR50 response (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

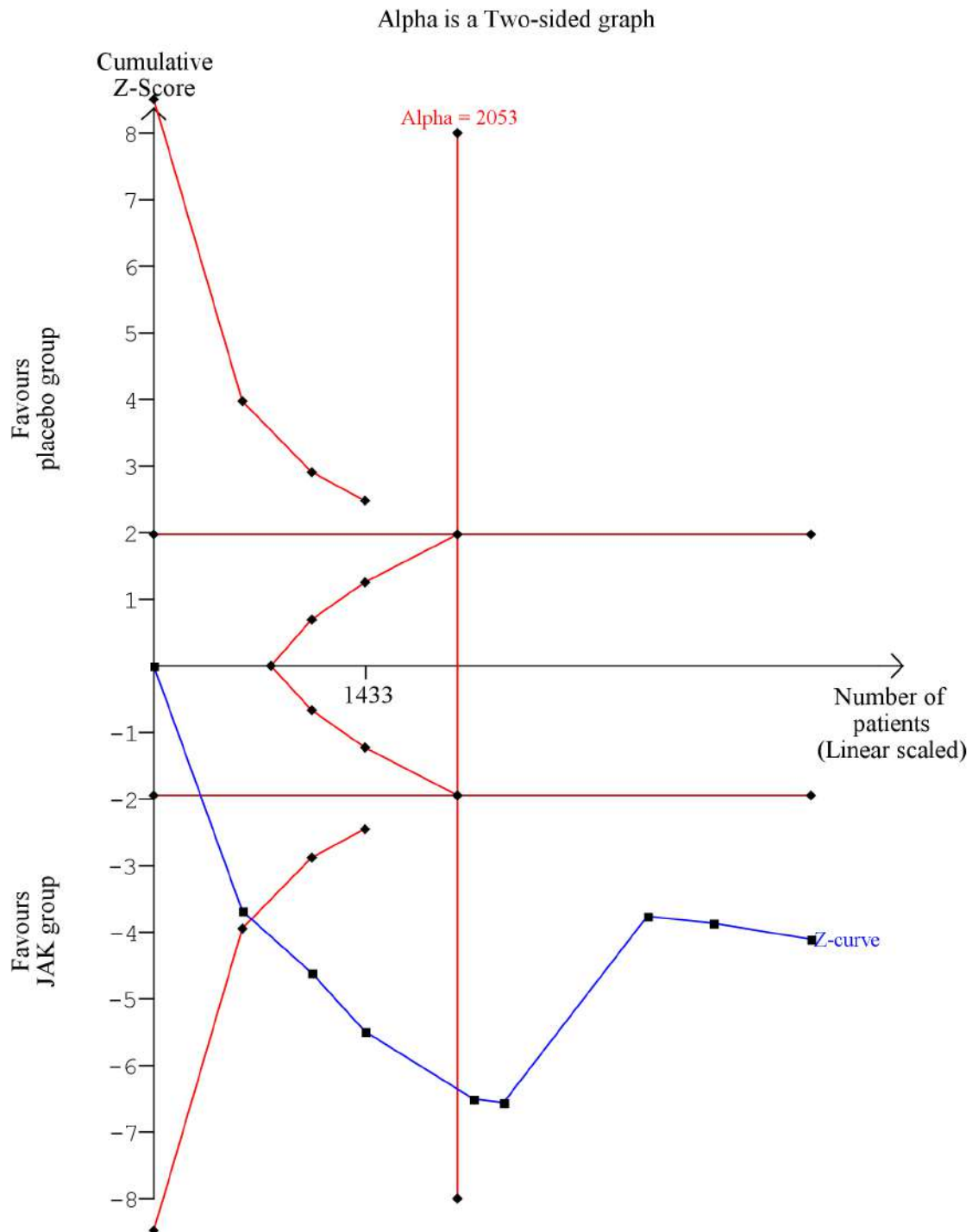**S108. TSA of ACR70 response (JAK inhibitors vs placebo)**

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

S109. TSA of HAQ-DI improvement (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph

S110. TSA of CDAI remission (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
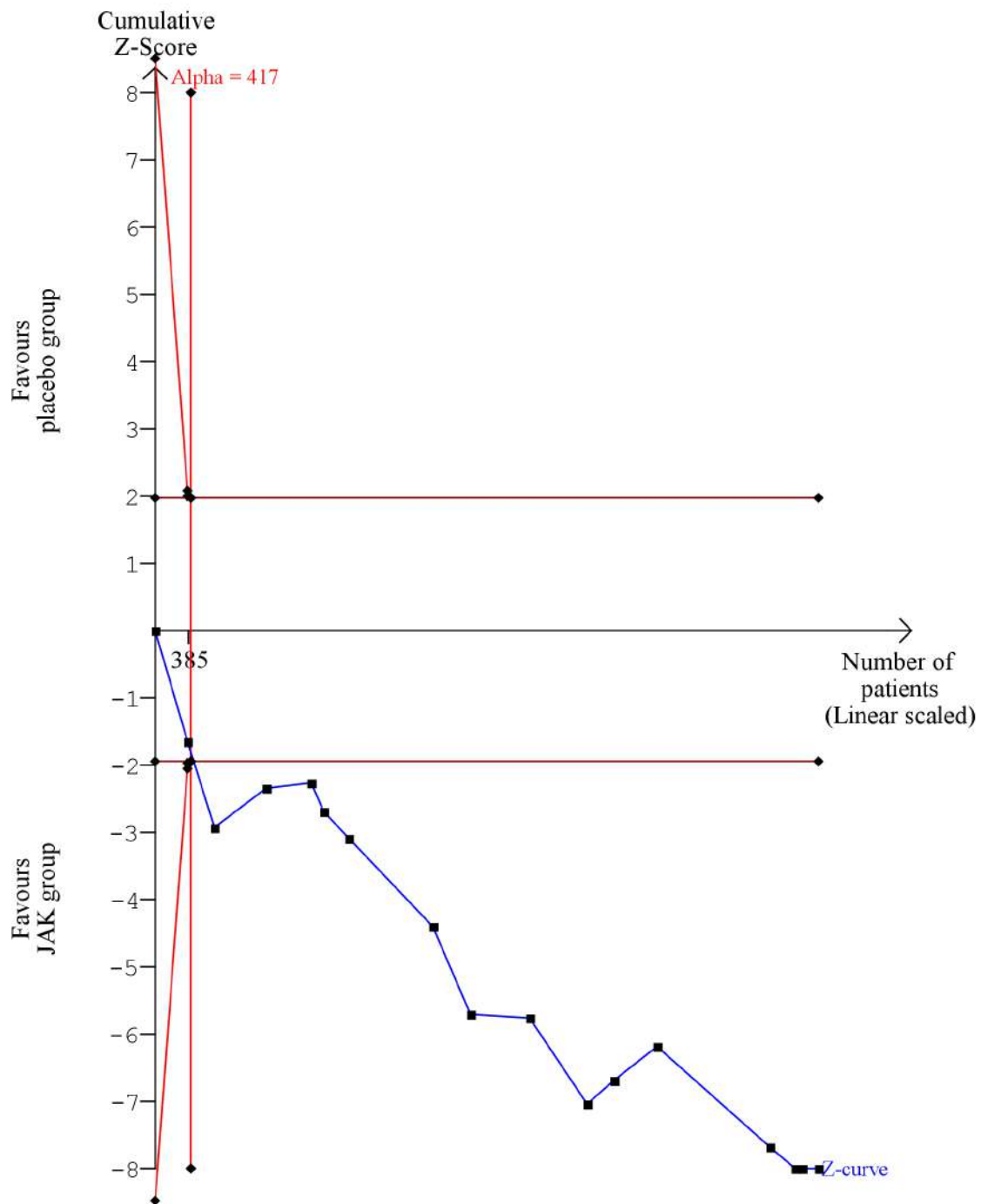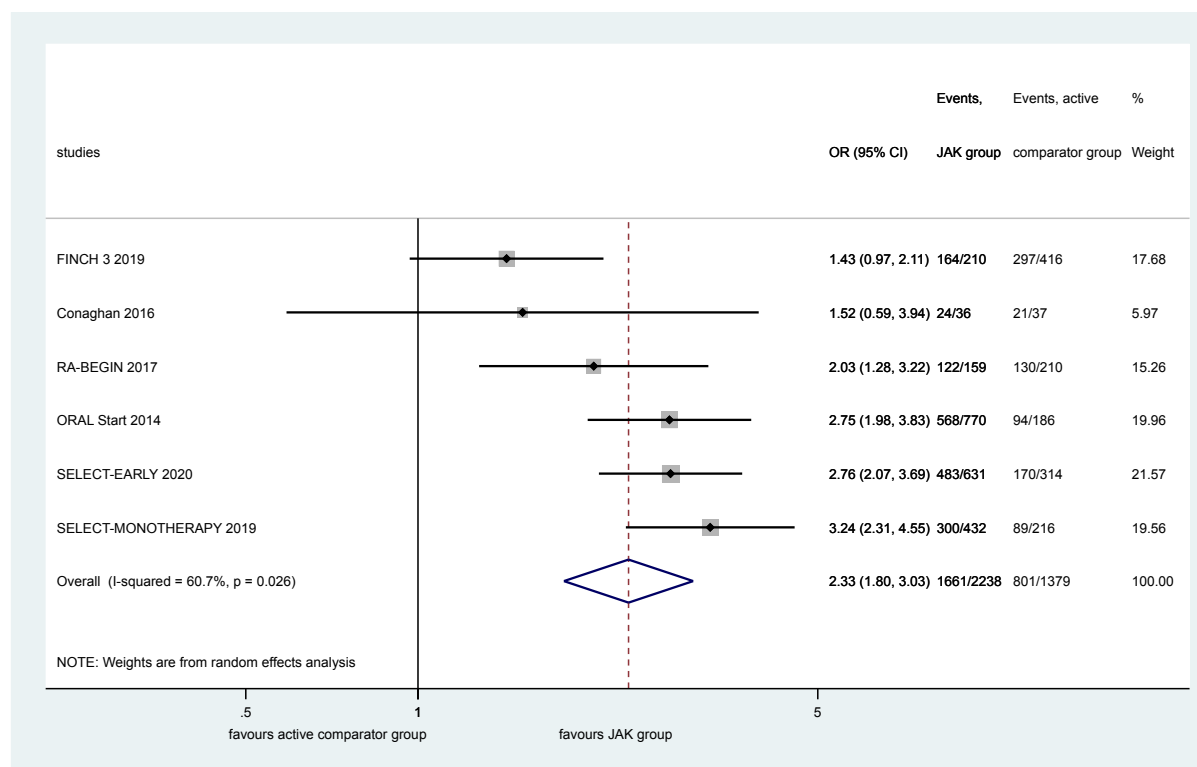
S111. TSA of SDAI remission (JAK inhibitors vs placebo)

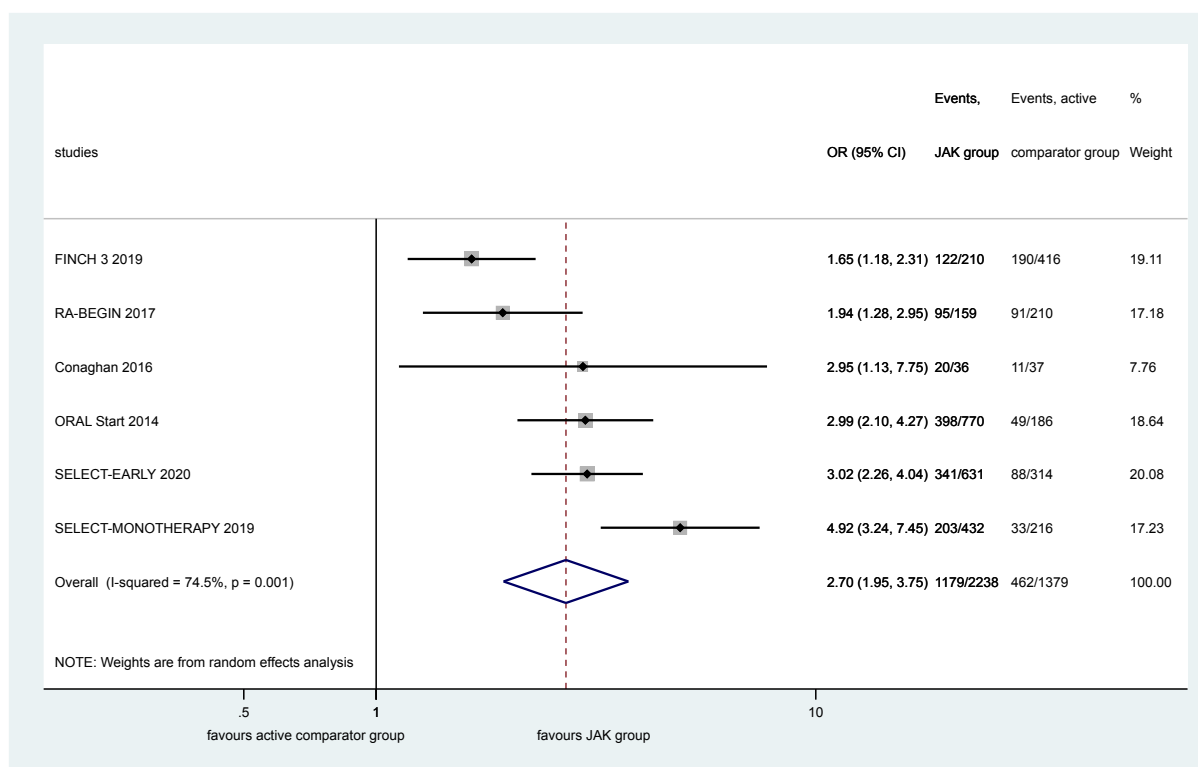The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

S112. TSA of DAS28-CRP remission (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph
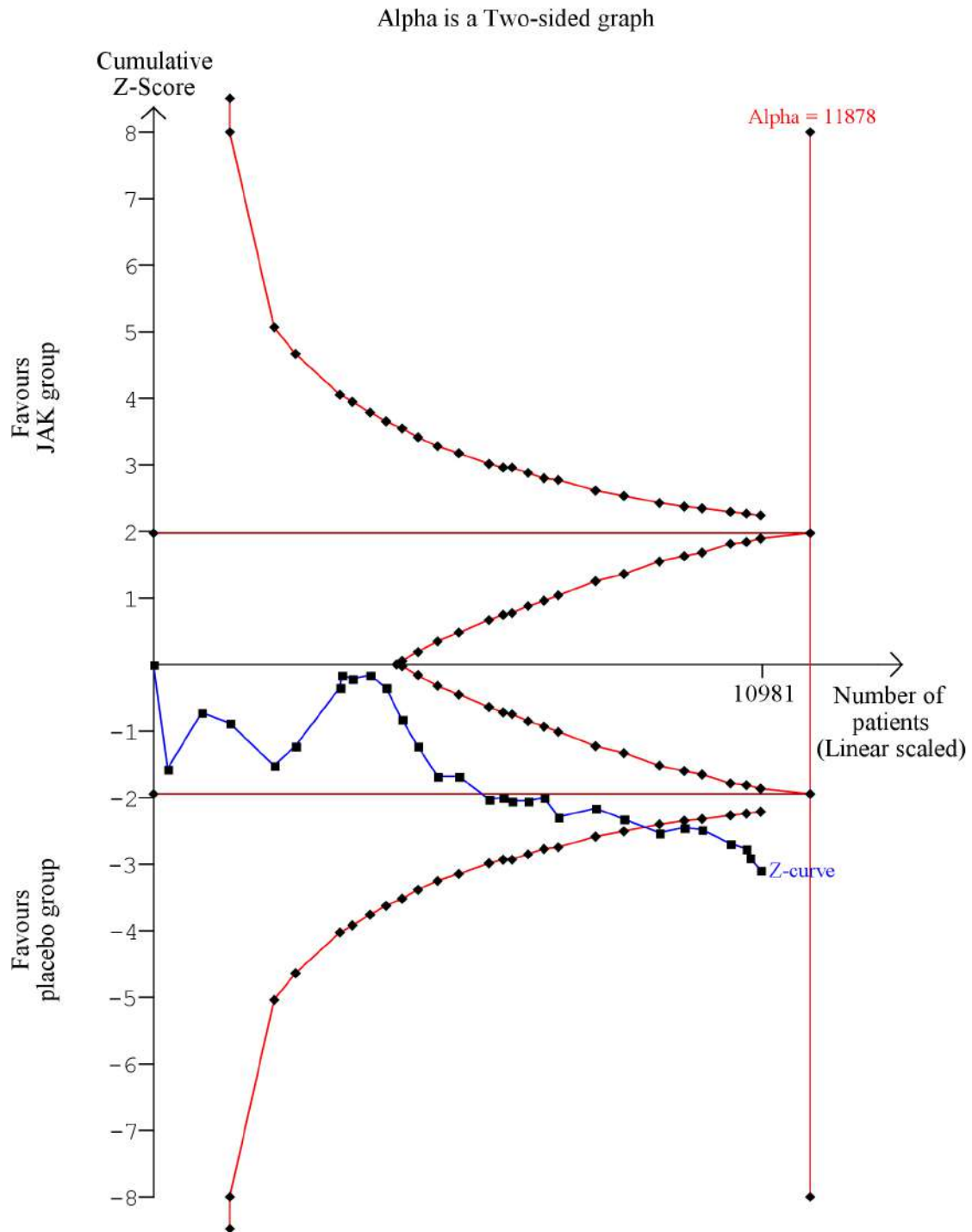
S113. TSA of side effects (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) but does not reach the line of required sample size, therefore showing the statistically significant superiority of placebo in this outcome but the addition of further studies might influence these results.

Several studies ((2015)Boyle (2020)Robinson) could not be included due to low information content.

Alpha is a Two-sided graph
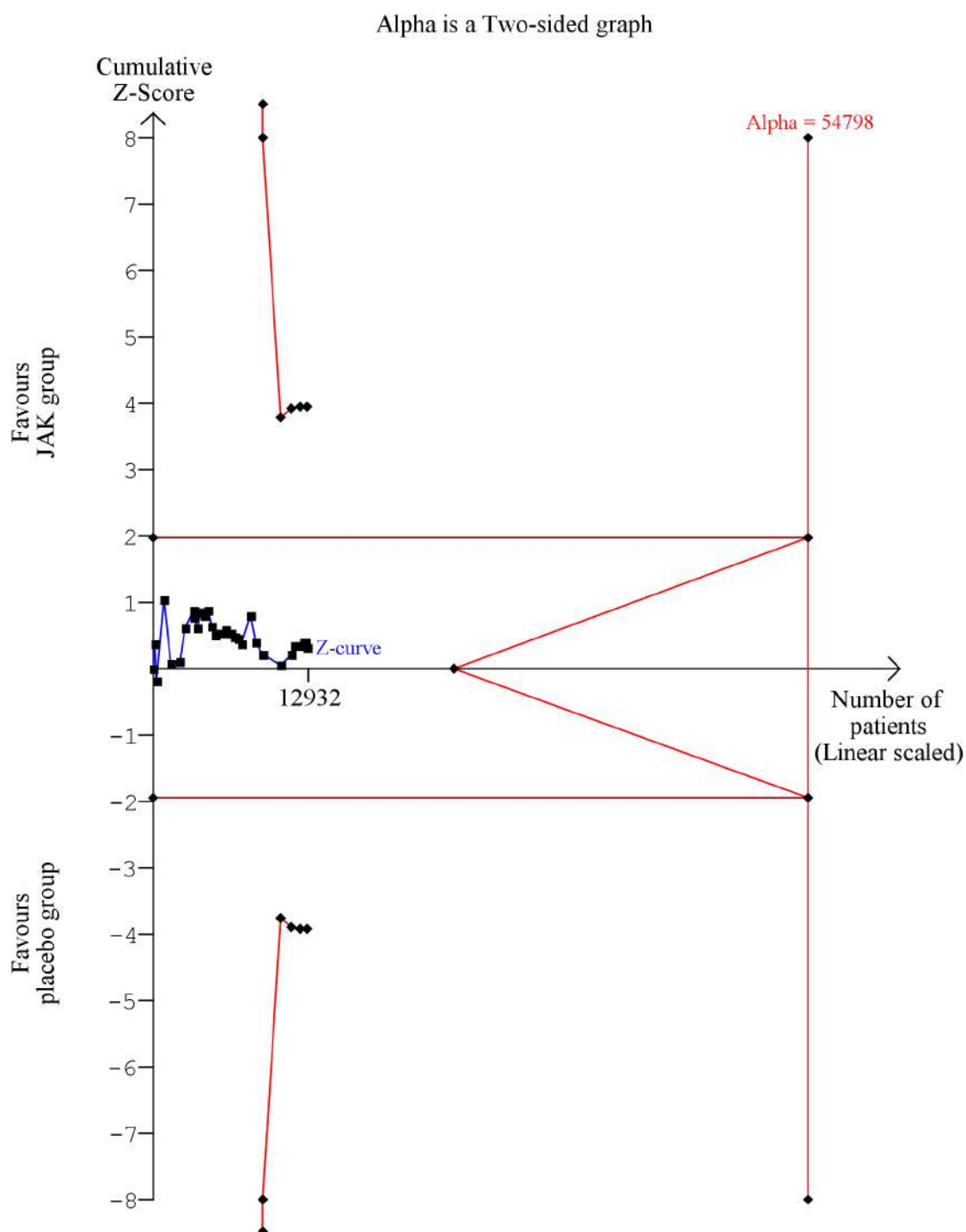
S114. TSA of serious side effects (JAK inhibitors vs placebo)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant

difference between groups in this outcome and the addition of further studies might influence these results. Several studies (2015)I4V-MC-JADA (2016)BALANCE I (2016)Genovese (2016)RA-BEACON (2016)Tanaka (2017)DARWIN2 (2017)RA-BEGIN (2018)SELECT-BEYOND (2019)RAJ3 (2020)RA-BALANCE (2020)Robinson) could not be included due to low information content.
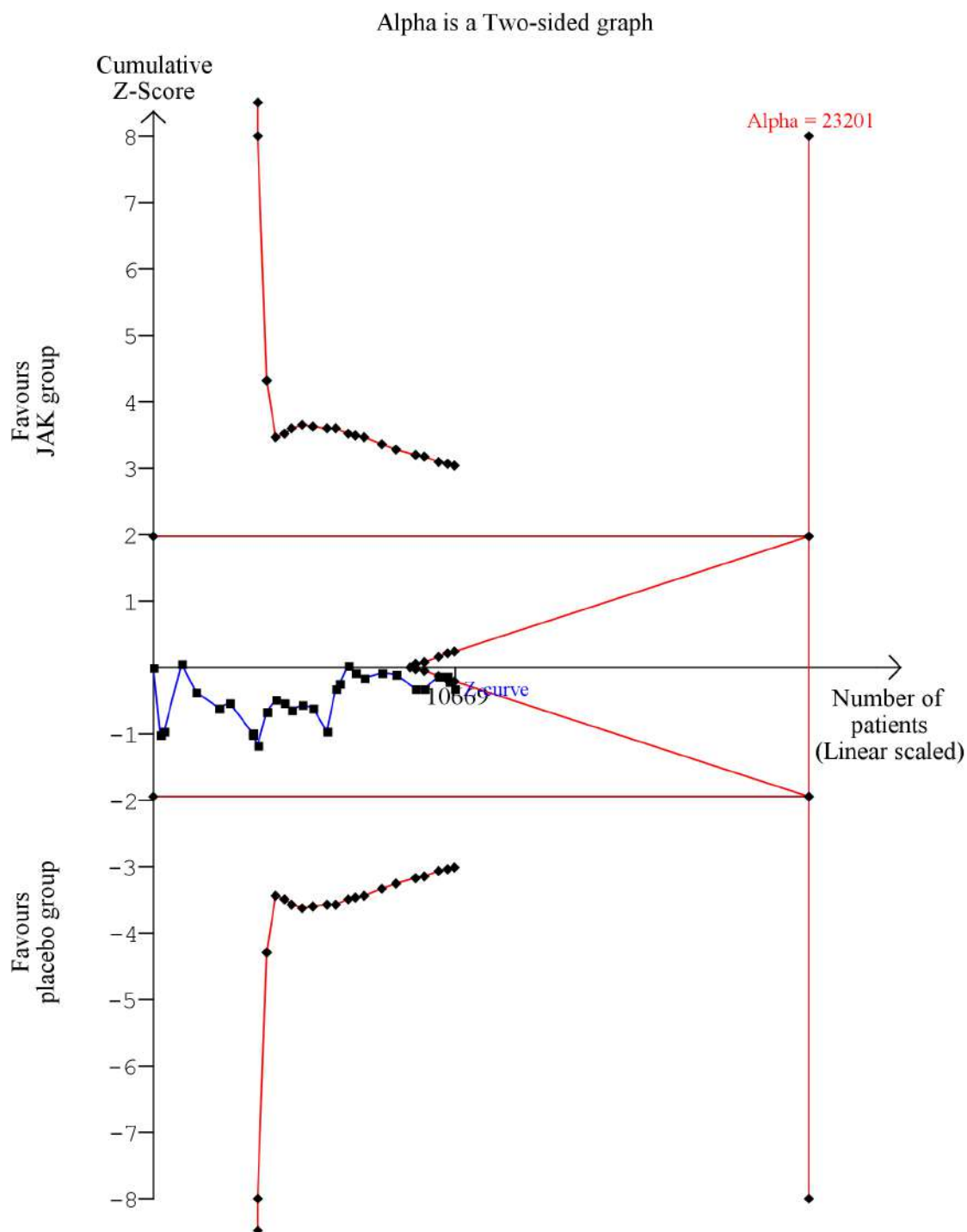
S115. TSA of discontinuation (JAK inhibitors vs placebo)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results. Several studies ((2011)Tanaka, (2015)Boyle, (2016)Tanaka, (2020)Robinson) could not be included due to low information content.
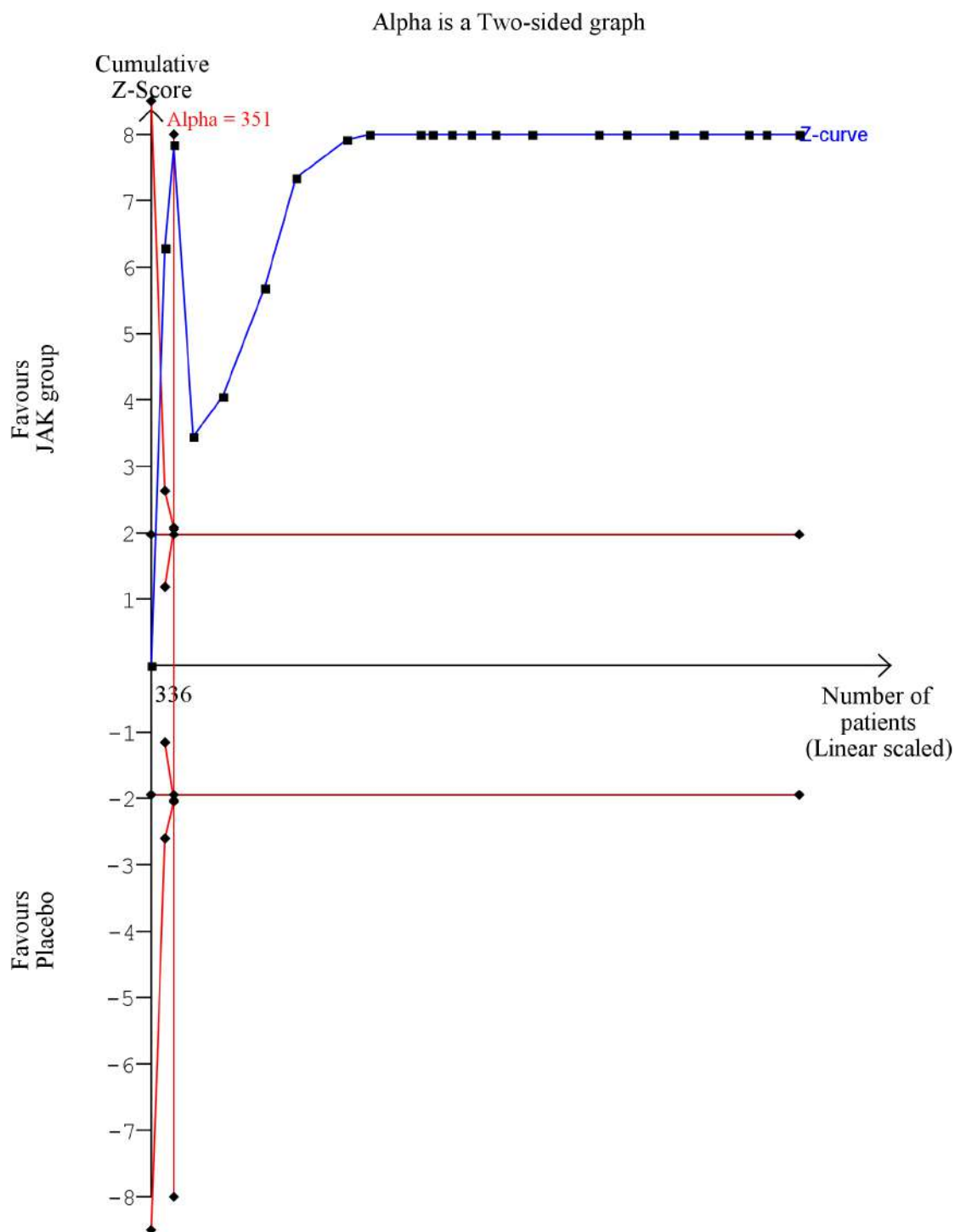
S116. TSA of Pain(VAS) (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
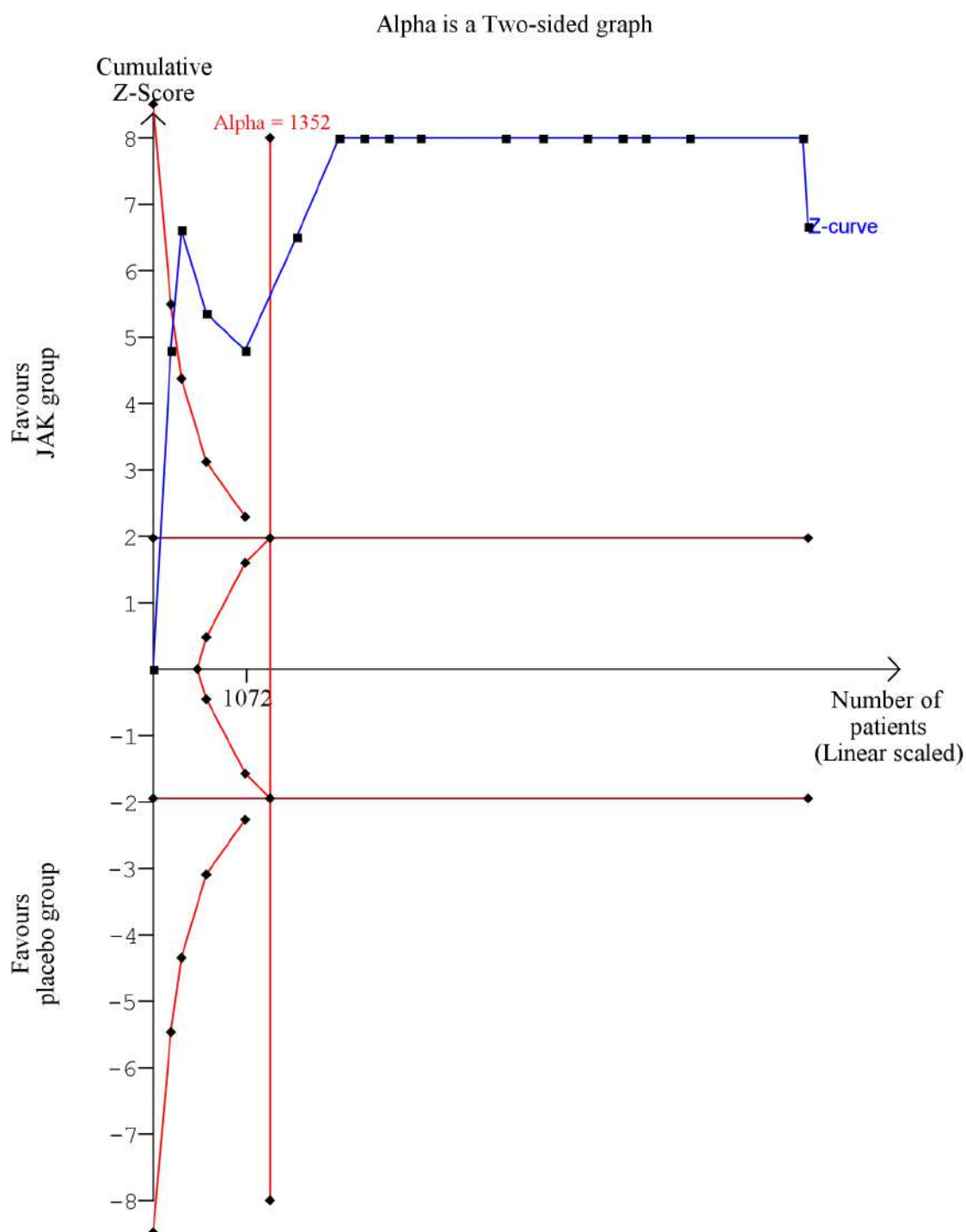
S117. TSA of PGA(VAS) (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
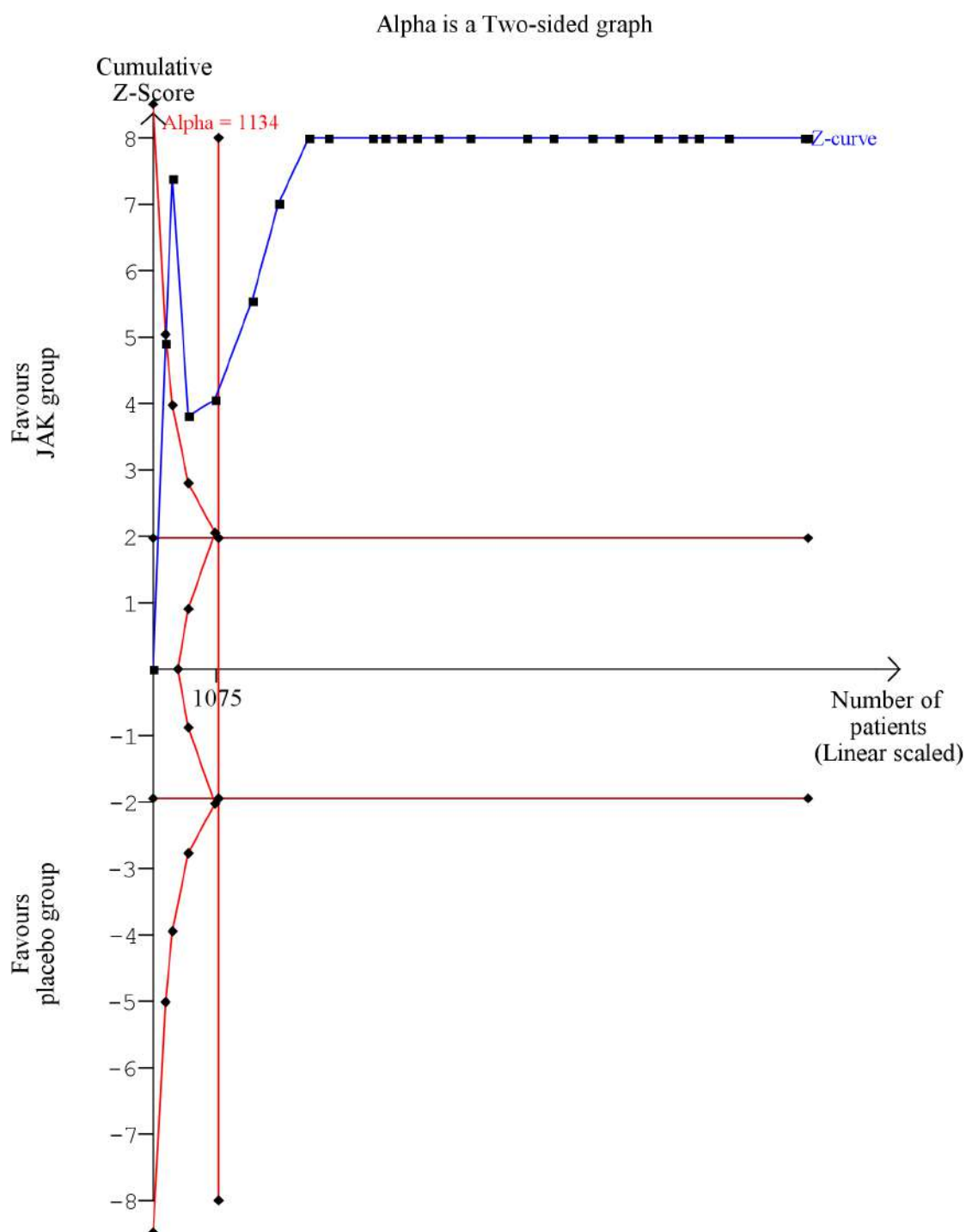
S118. TSA of PtGA(VAS) (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
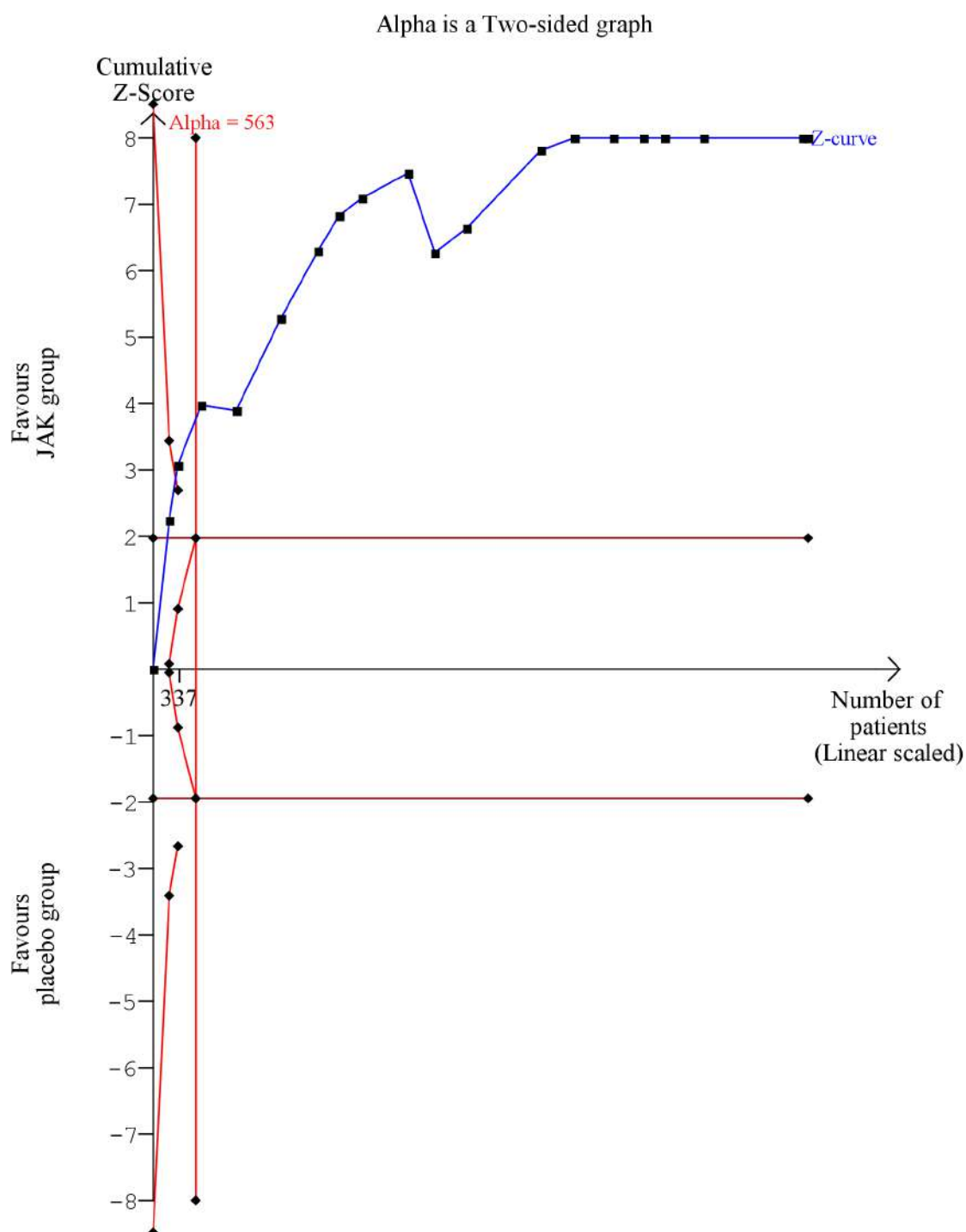
S119. TSA of Swollen Joint Count (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
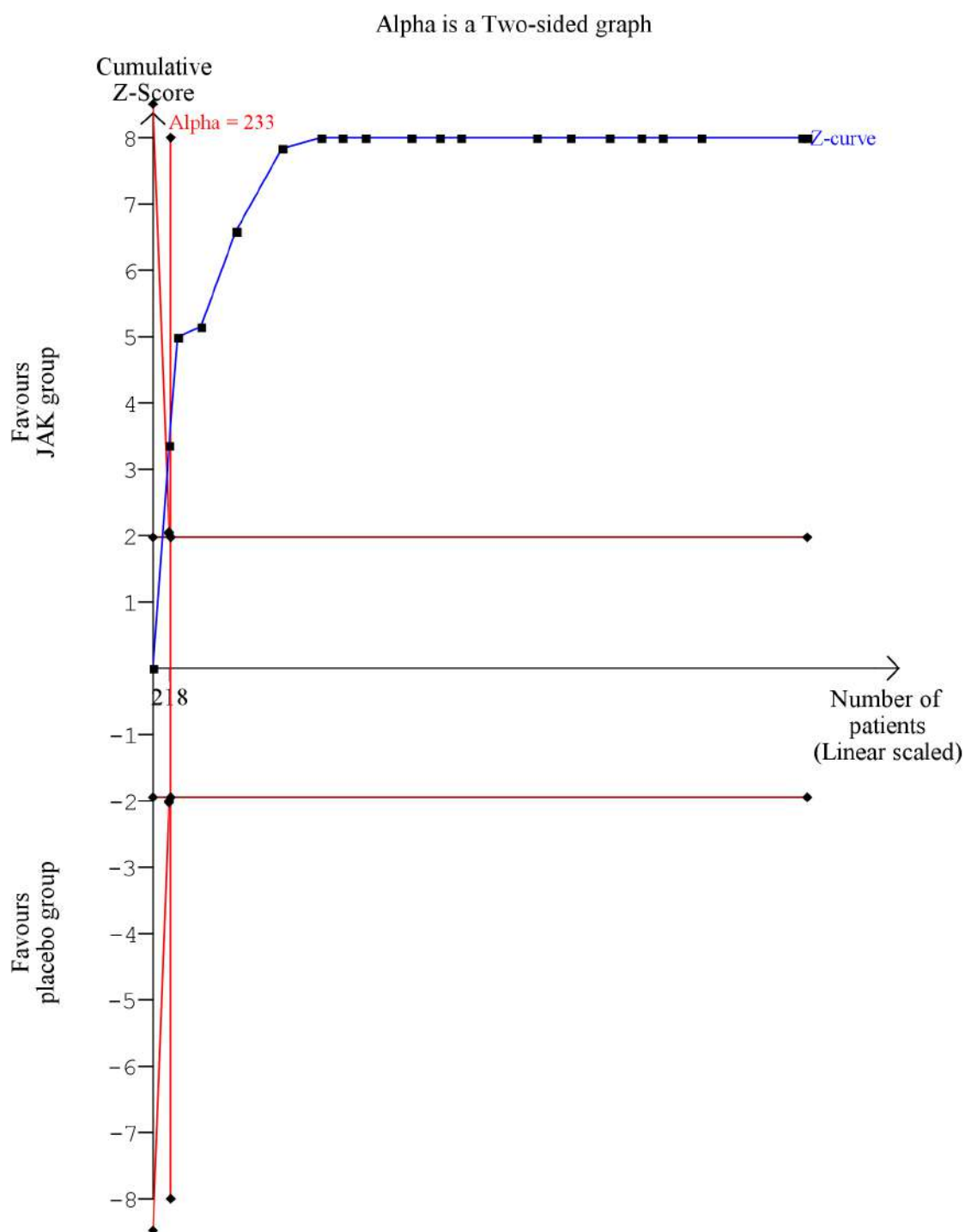
S120. TSA of Tender Joint Count (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
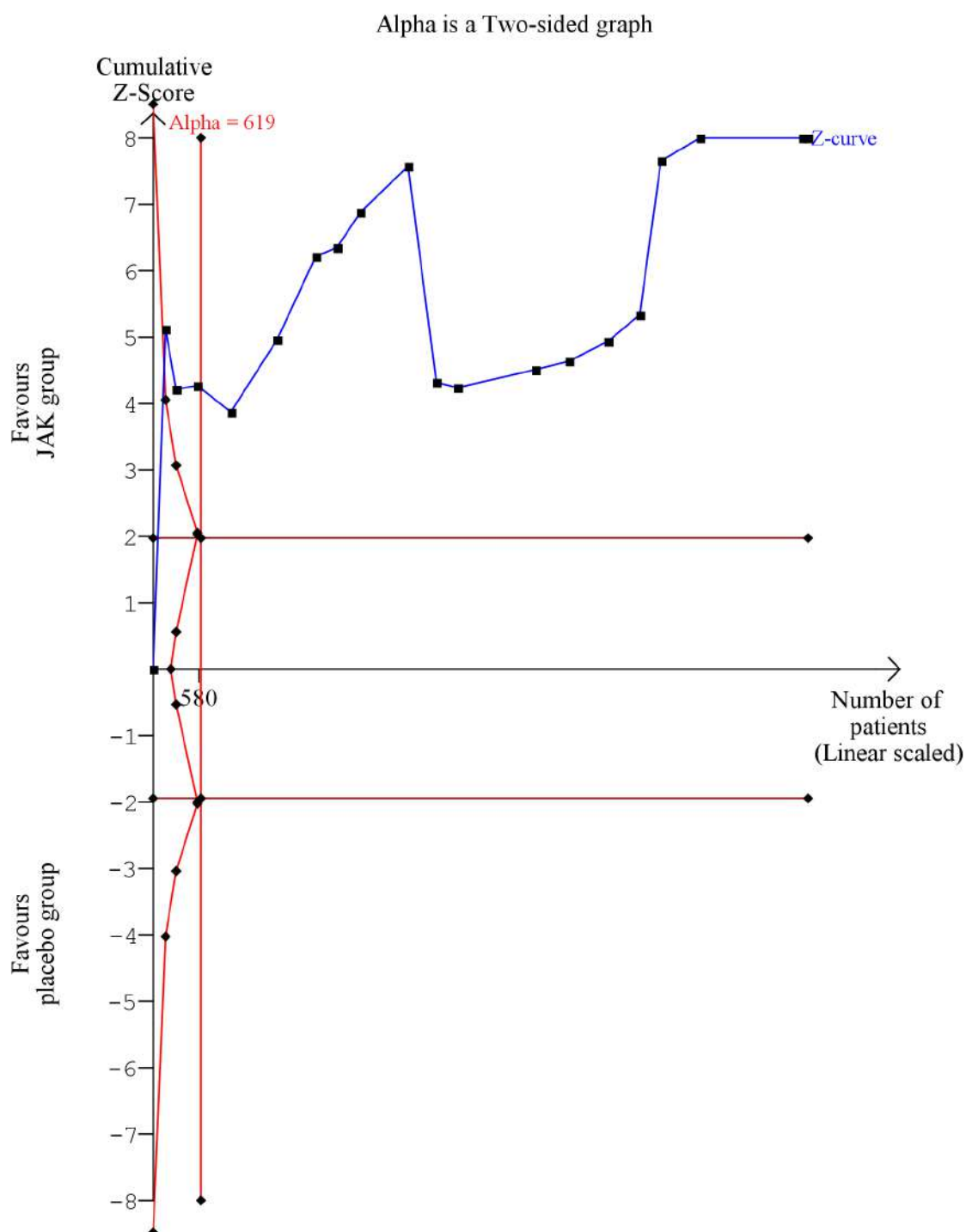
S121. TSA of CRP (mg/l) (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
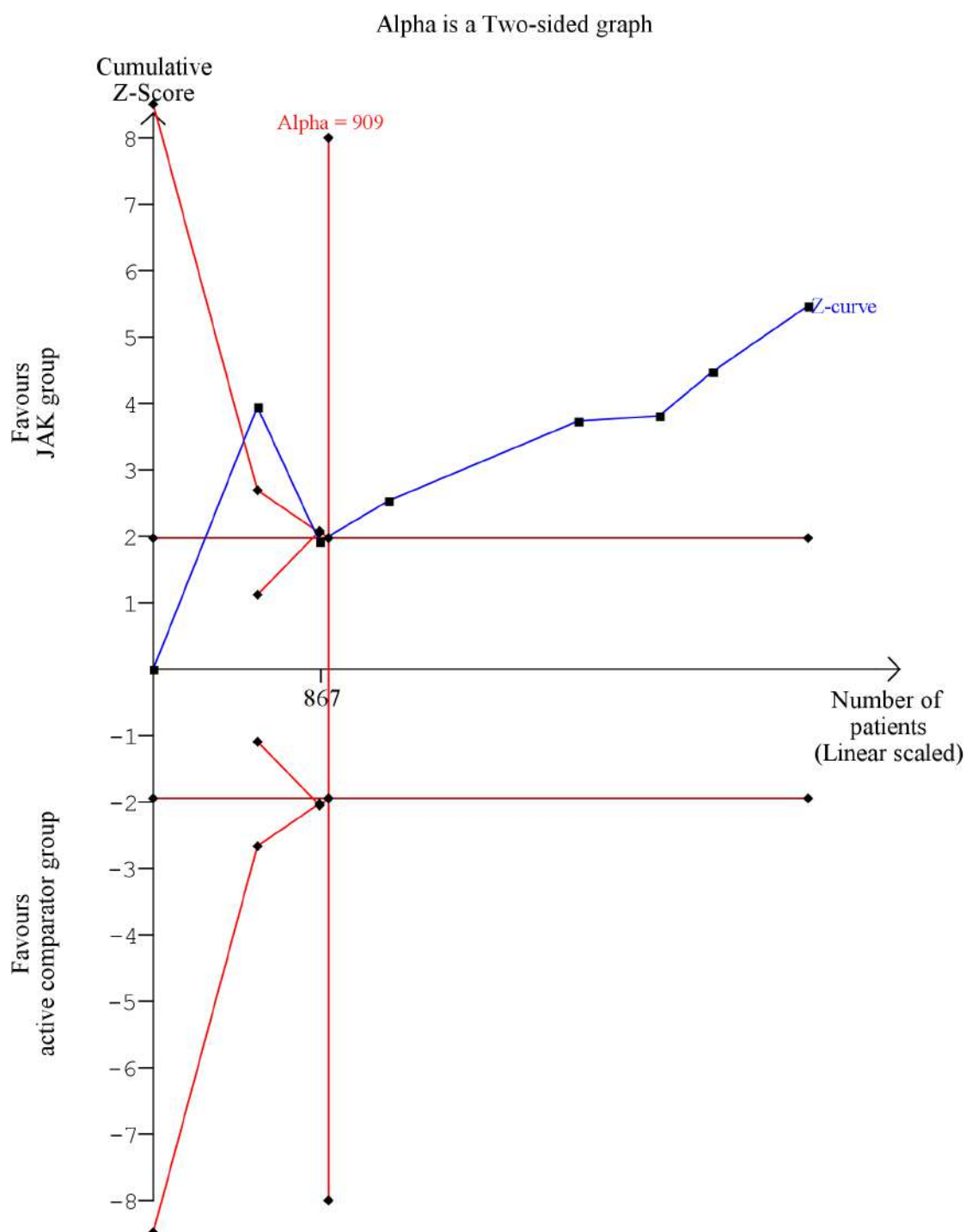
S122. TSA of ESR (mm/hour) (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
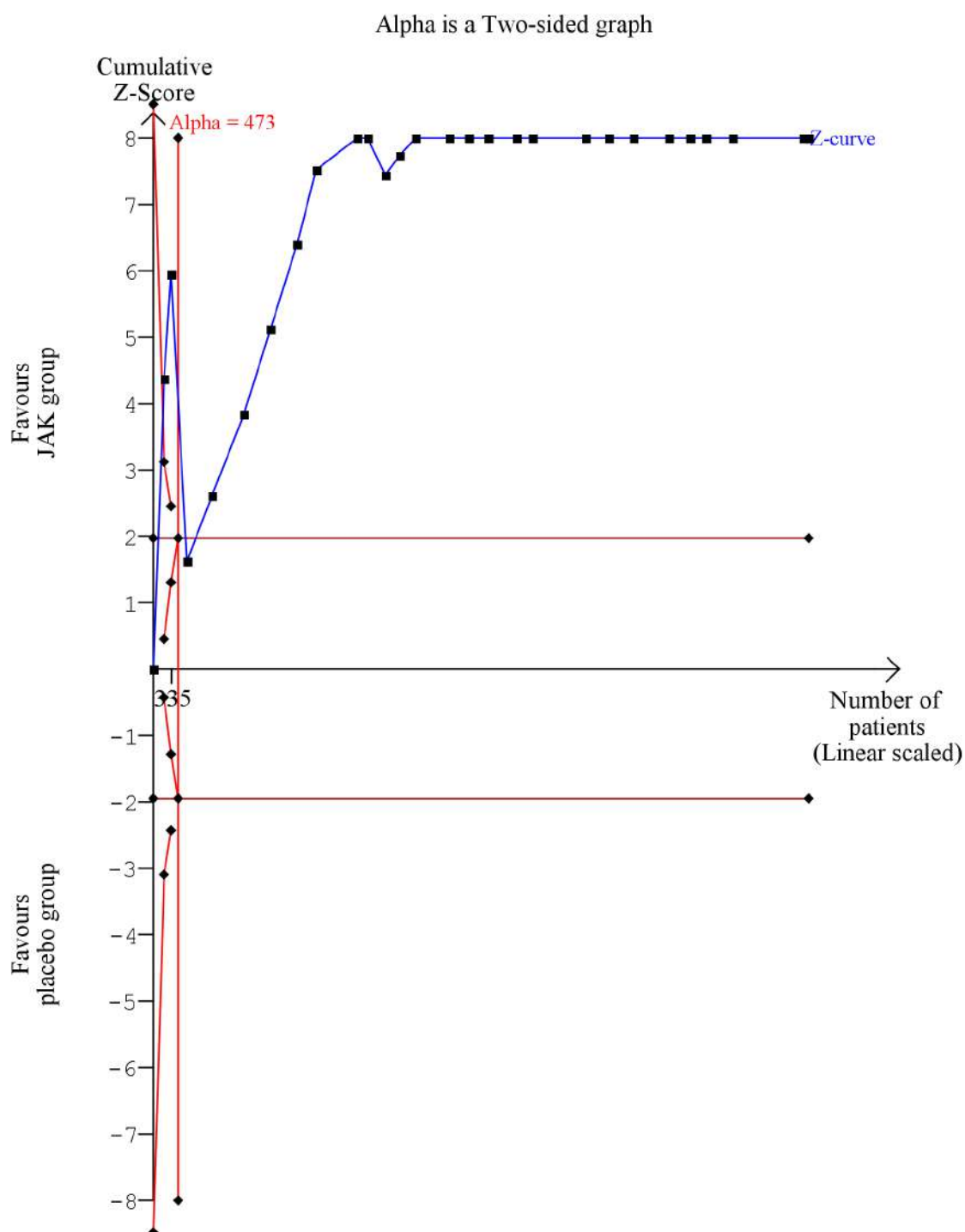
S123. TSA of HAQ-DI difference (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
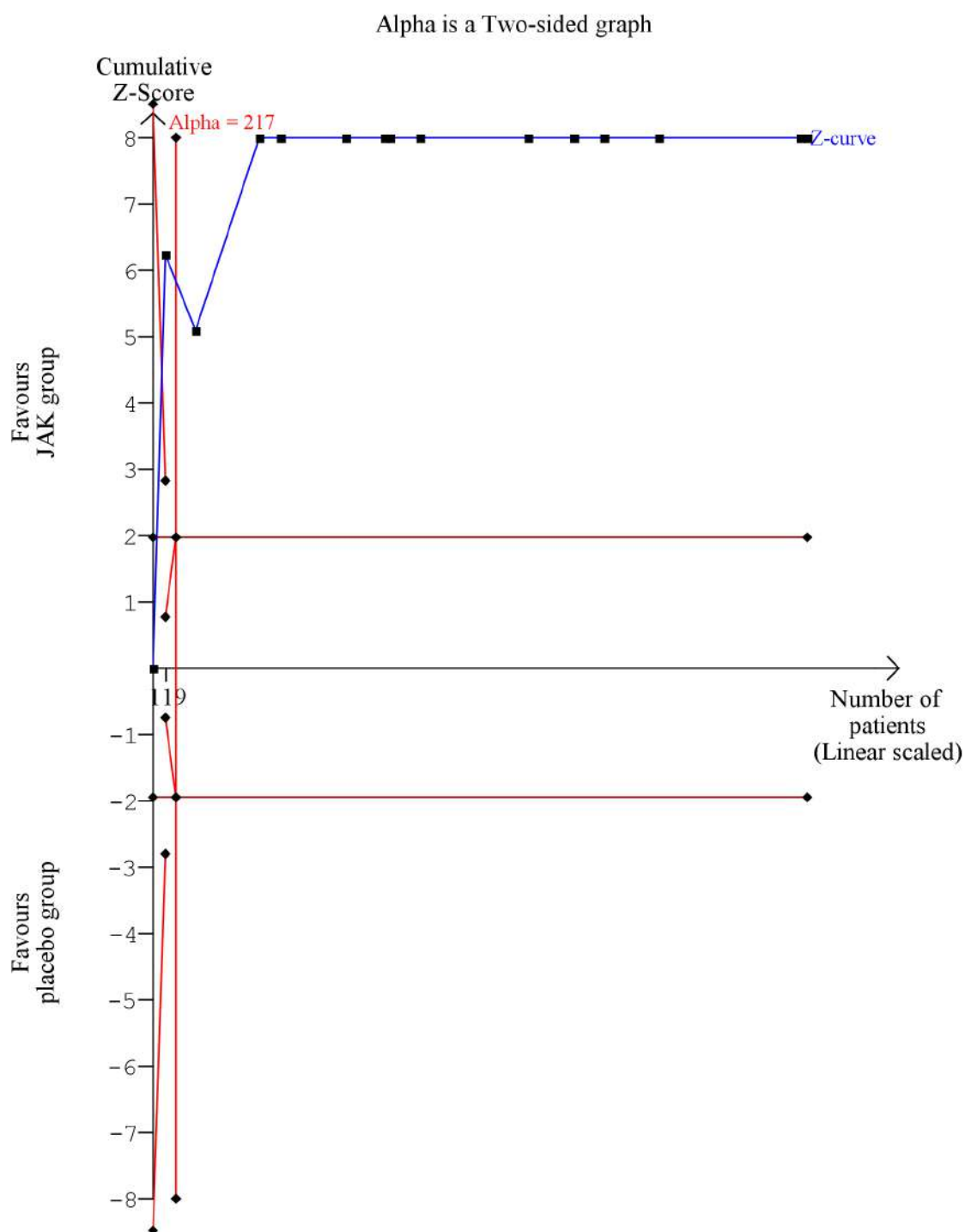
S124. TSA of DAS28-CRP difference (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.
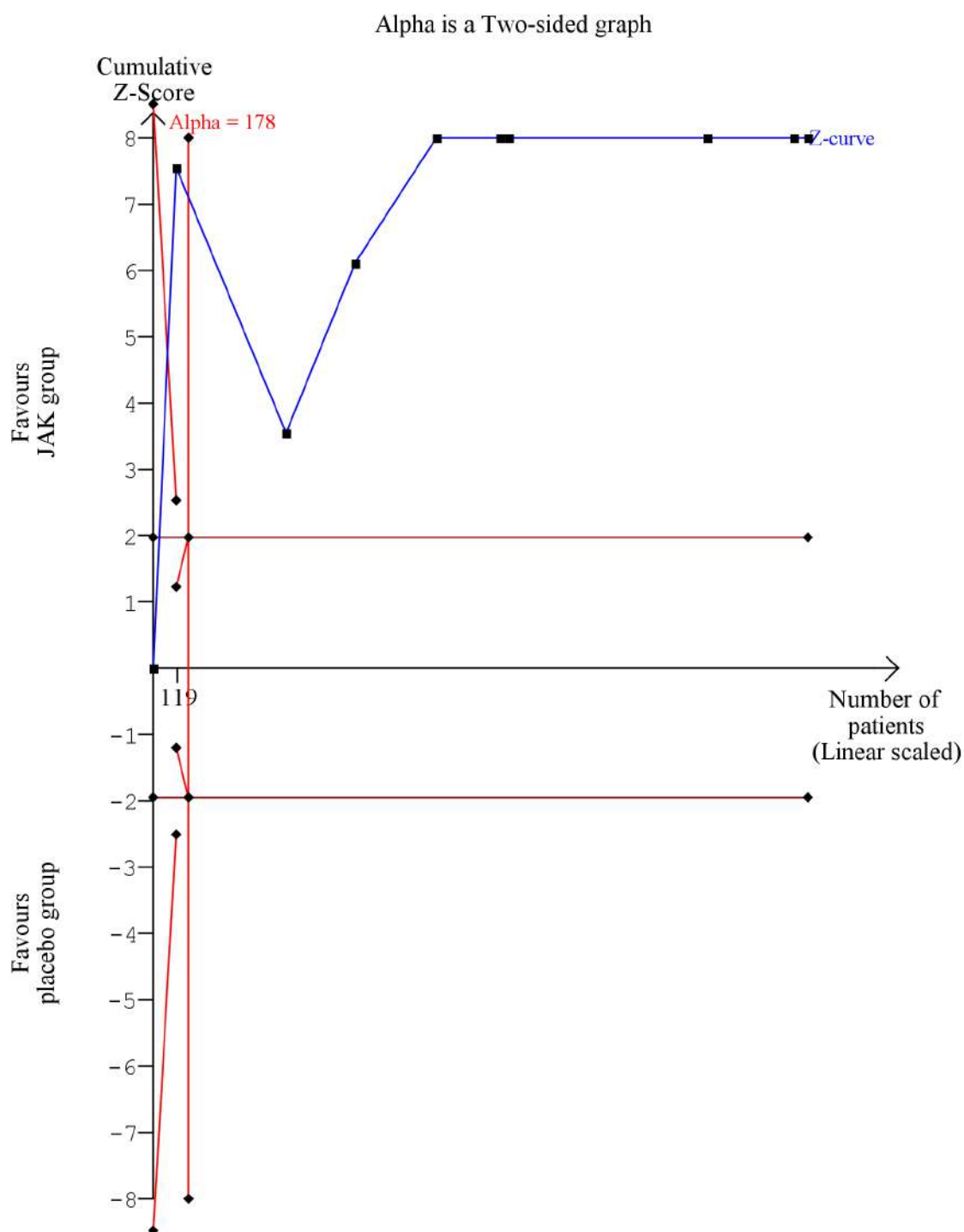
S125. TSA of DAS28-ESR difference (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.
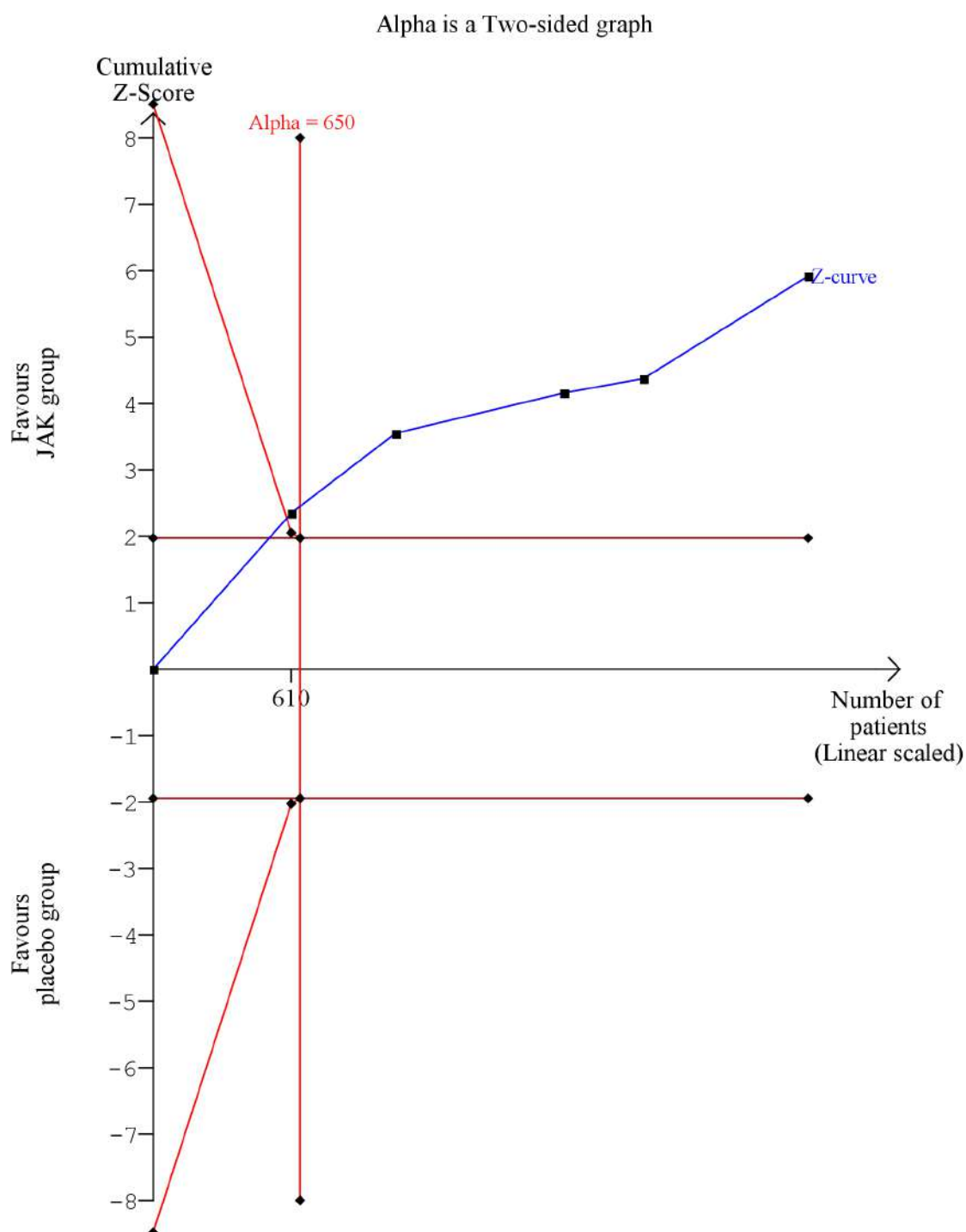
S126. TSA of MOS-Sleep (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.
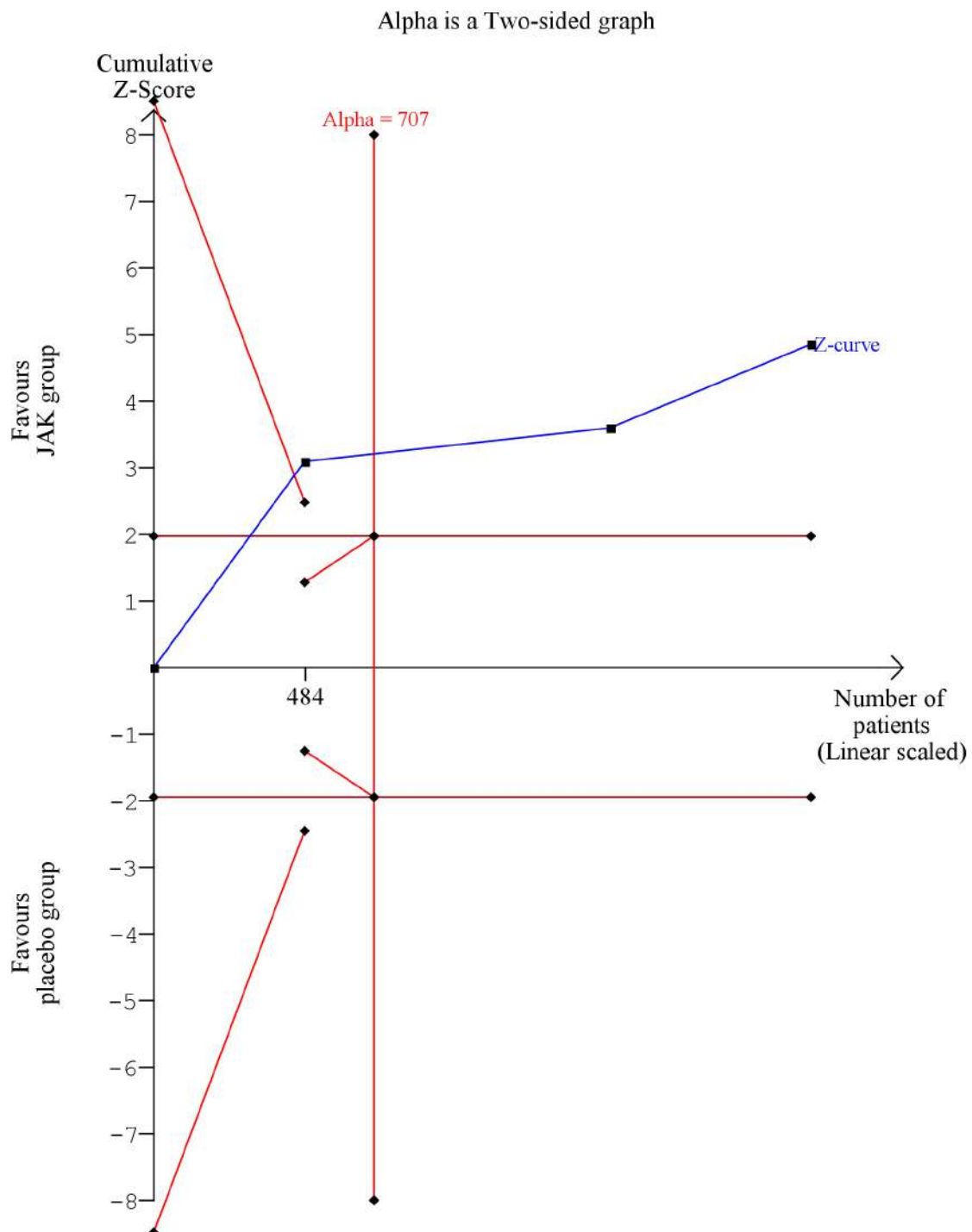
S127. TSA of WPAI AI (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.
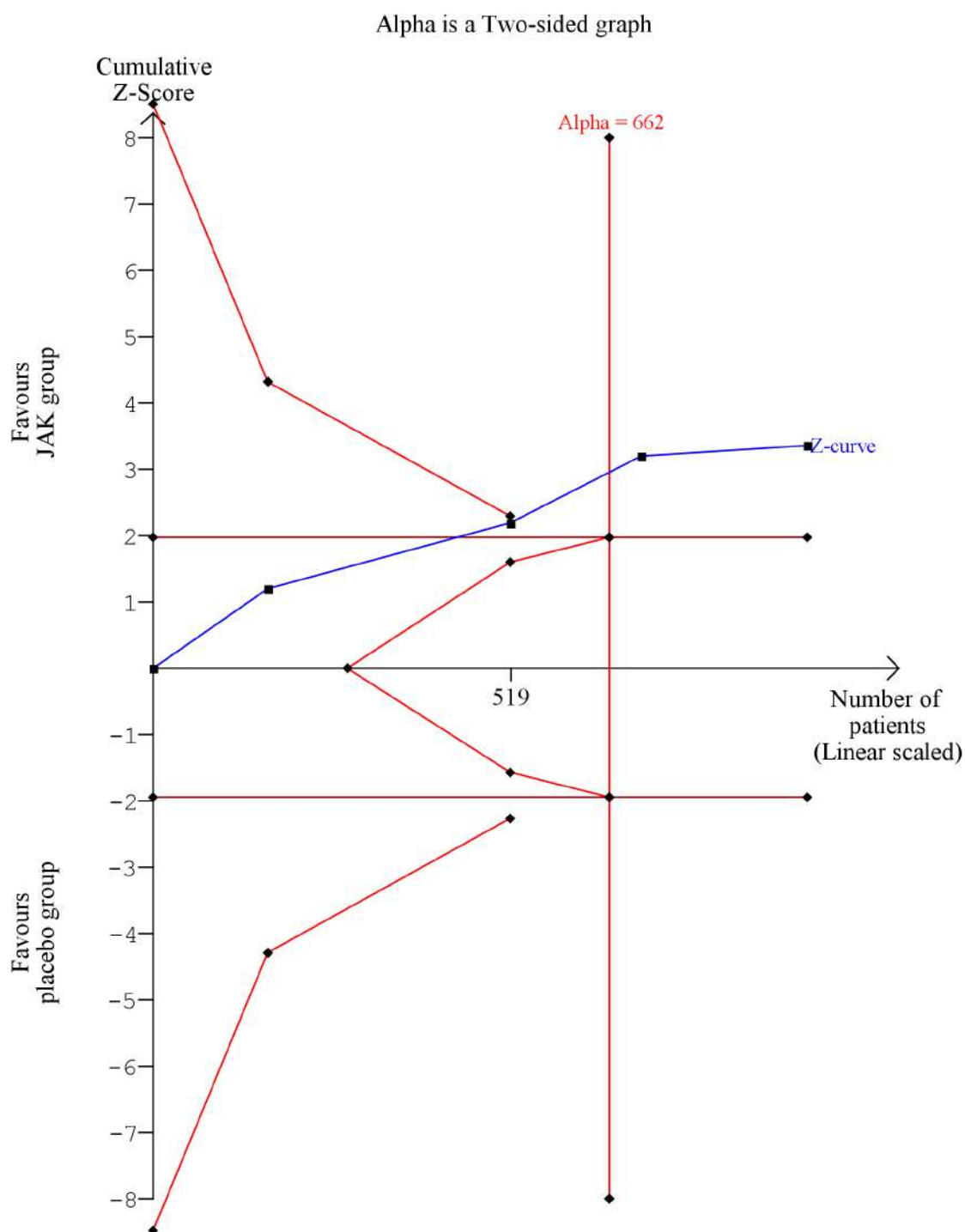
S128. TSA of WPAI A (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.
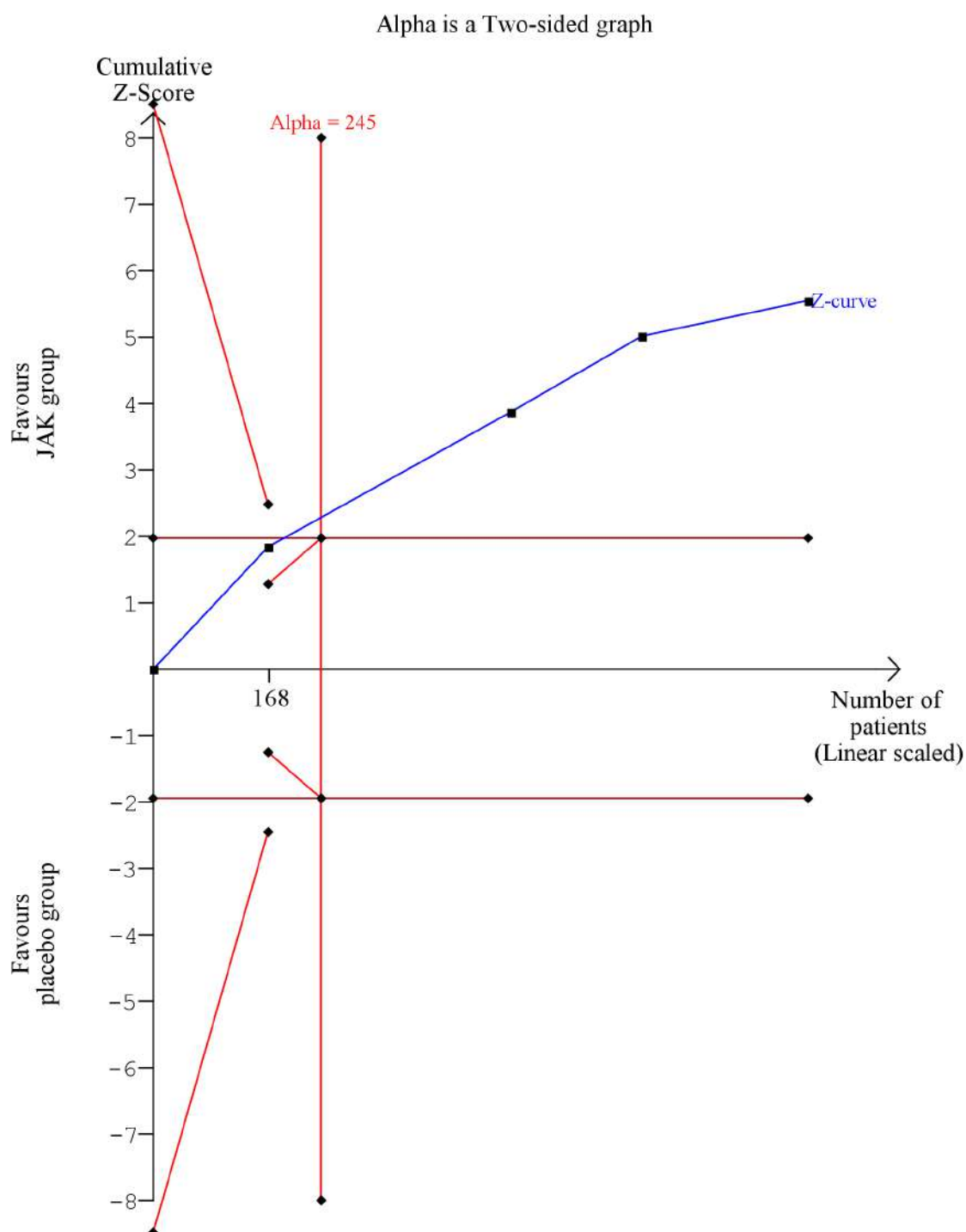
S129. TSA of WPAI OWI (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

S130. TSA of WPAI P (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph

S131. TSA of MJS duration (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size,

therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph

S132. TSA of SF-36 MCS (JAK inhibitors vs placebo)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size,

therefore showing the statistically significant superiority of JAK inhibitors in outcome and the addition of further studies would not influence these results.

(2) <u>JAK inhibitors in monotherapy compared to MTX</u>

studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
FINCH 3 2019	1.43 (0.97, 2.11)	164/210	297/416	17.68
Conaghan 2016	1.52 (0.59, 3.94)	24/36	21/37	5.97
RA-BEGIN 2017	2.03 (1.28, 3.22)	122/159	130/210	15.26
ORAL Start 2014	2.75 (1.98, 3.83)	568/770	94/186	19.96
SELECT-EARLY 2020	2.76 (2.07, 3.69)	483/631	170/314	21.57
SELECT-MONOTHERAPY 2019	3.24 (2.31, 4.55)	300/432	89/216	19.56
Overall (I-squared = 60.7%, p = 0.026)	2.33 (1.80, 3.03)	1661/2238	801/1379	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S133. Forest plot of studies comparing the number of patients reaching at least 20% in ACR criteria between patients treated with JAK inhibitors and MTX within 6 months (ACR20 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
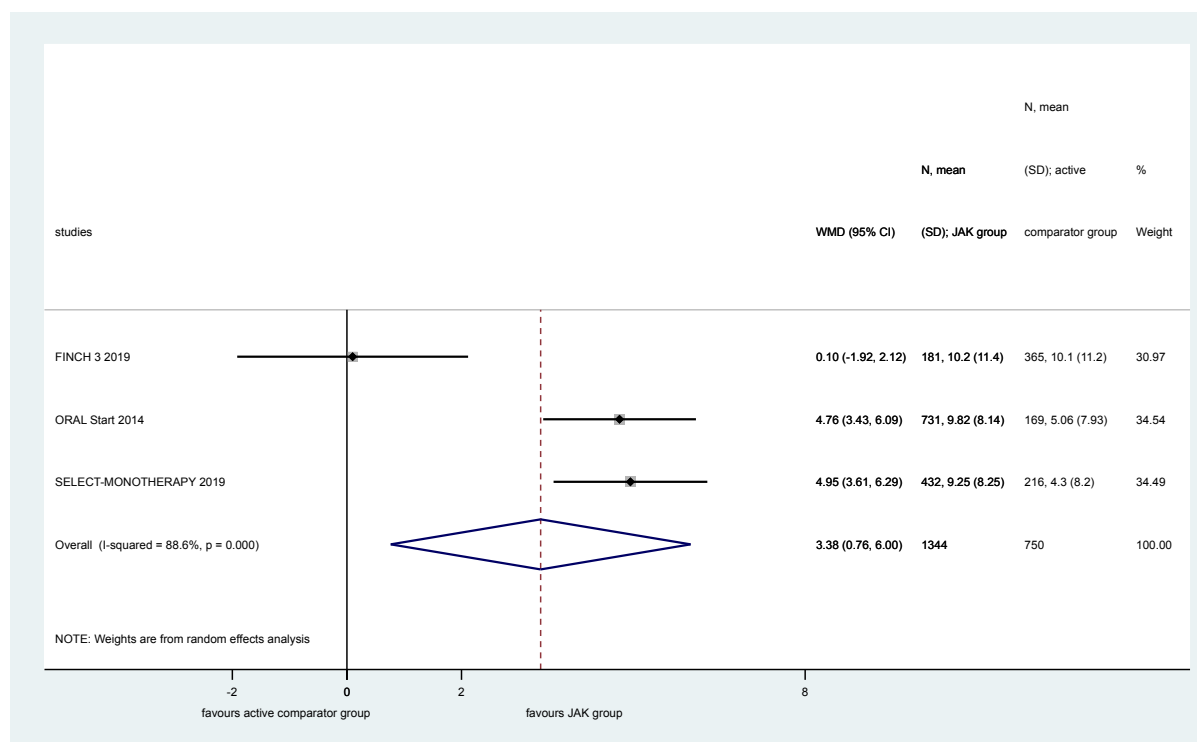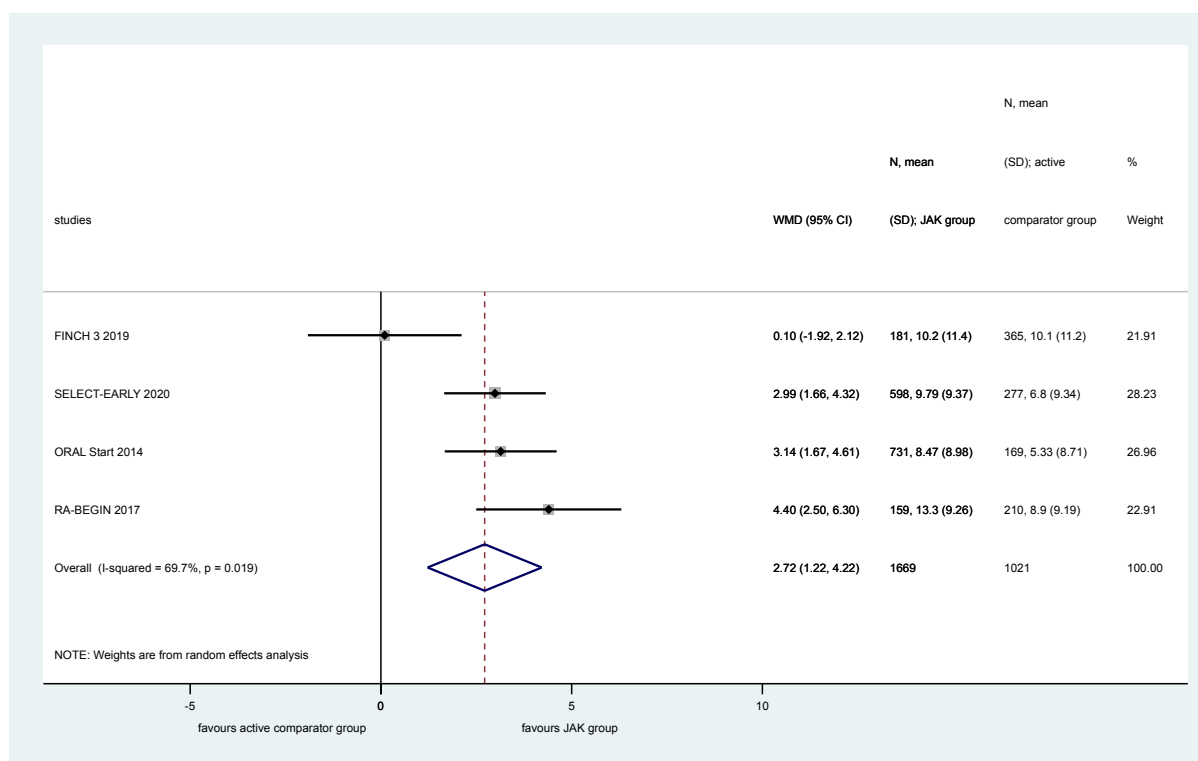
S134 a. Forest plot of studies comparing the number of patients reaching at least 50% in ACR criteria between patients treated with JAK inhibitors and MTX within 6 months (ACR50 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
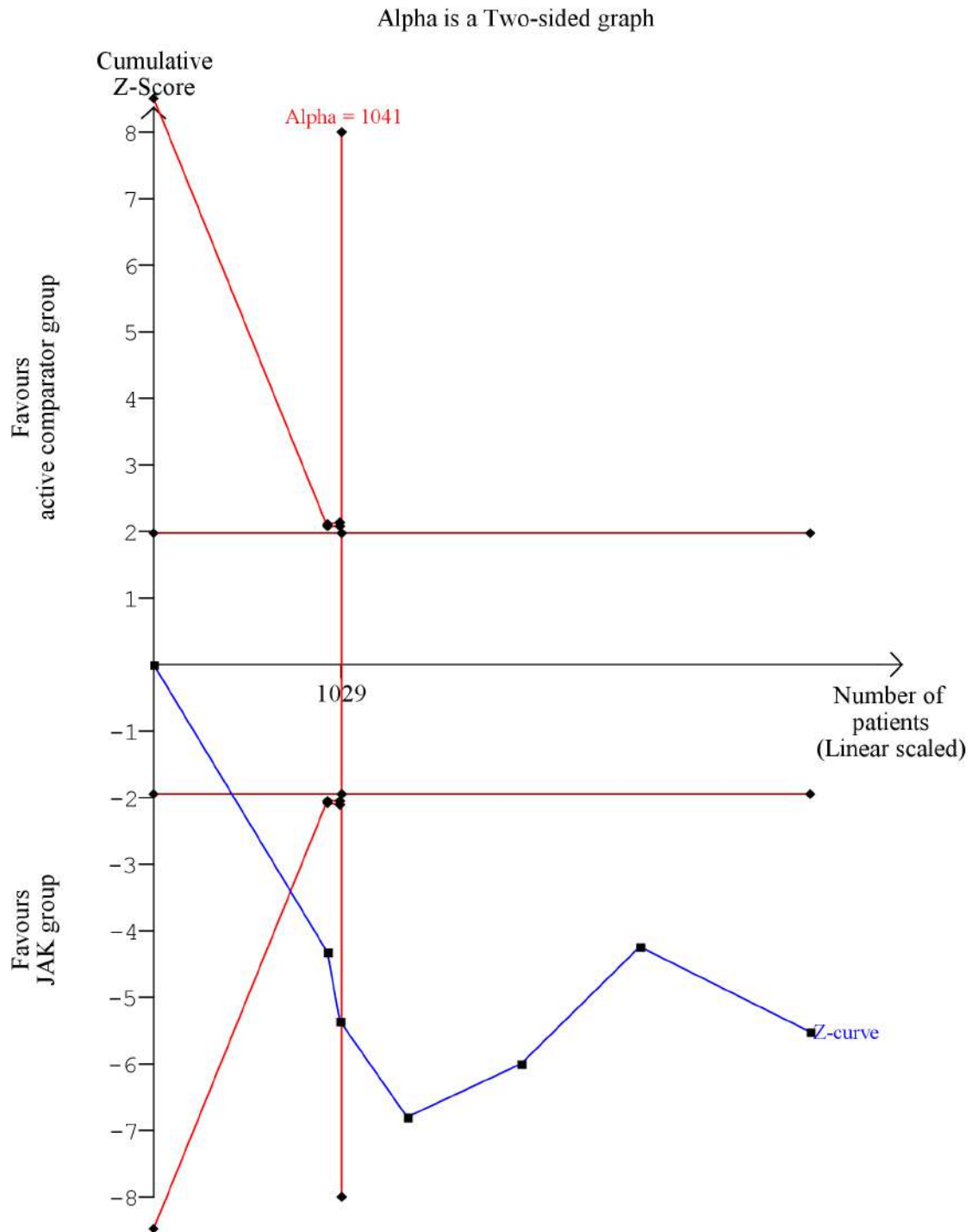
studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
FINCH 3 2019	1.65 (1.18, 2.31)	122/210	190/416	23.92
RA-BEGIN 2017	1.94 (1.28, 2.95)	95/159	91/210	20.14
Conaghan 2016	2.95 (1.13, 7.75)	20/36	11/37	6.93
ORAL Start 2014	2.99 (2.10, 4.27)	398/770	49/186	22.96
SELECT-EARLY 2020	3.02 (2.26, 4.04)	341/631	88/314	26.05
Overall (I-squared = 59.4%, p = 0.043)	2.38 (1.80, 3.16)	976/1806	429/1163	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group
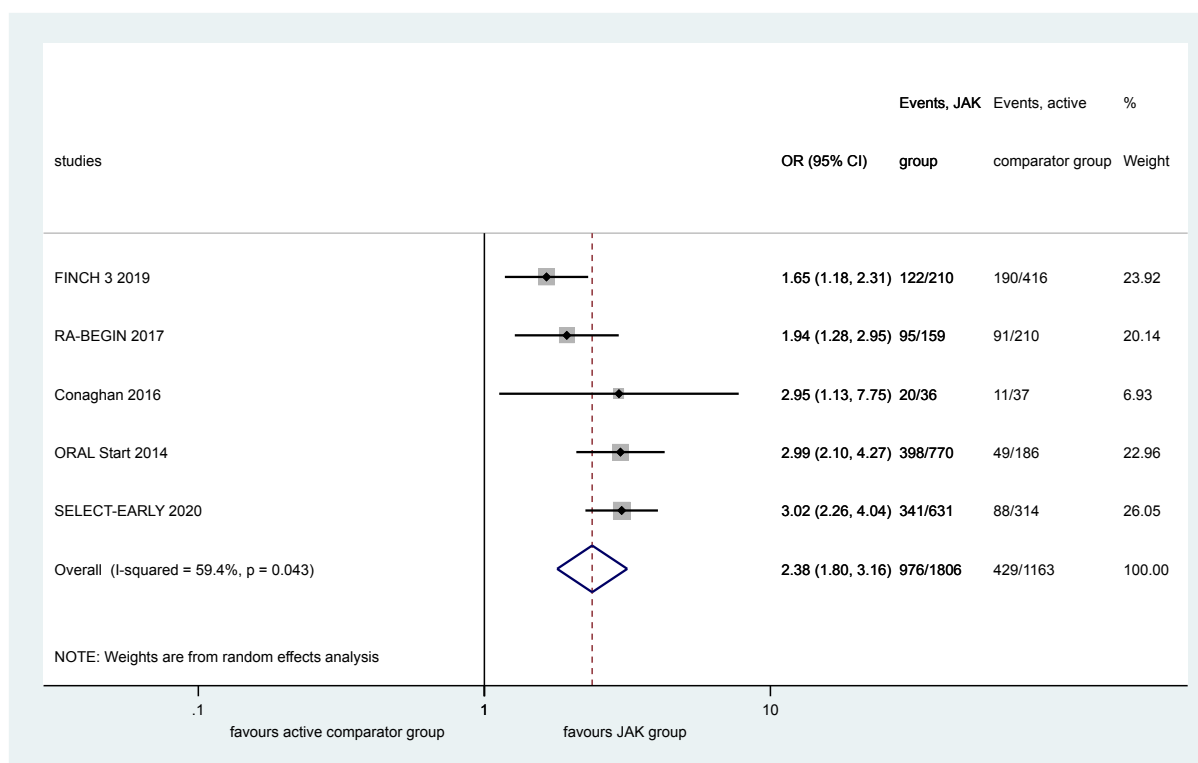
S134 b. Forest plot of studies comparing the number of patients reaching at least 50% in ACR criteria between patients treated with JAK inhibitors and MTX within 6 months (ACR50 response) – Secondary analysis only involving studies conducted in a MTX naïve population

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
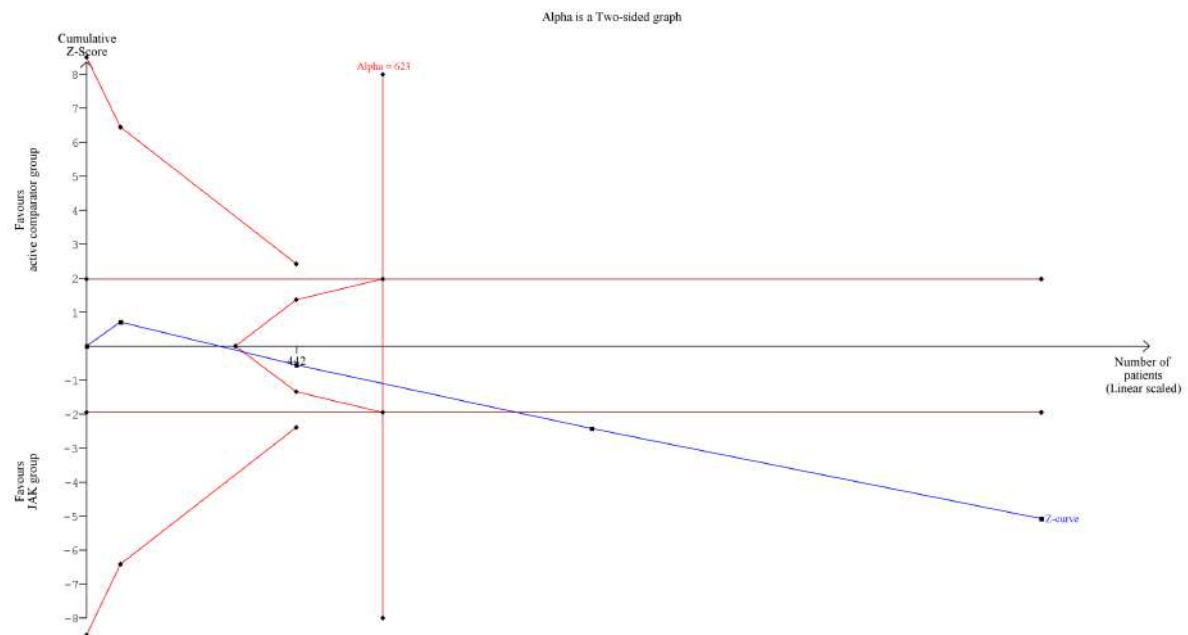
S135 a. Forest plot of studies comparing the number of patients reaching at least 70% in ACR criteria between patients treated with JAK inhibitors and MTX within 6 months (ACR70 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
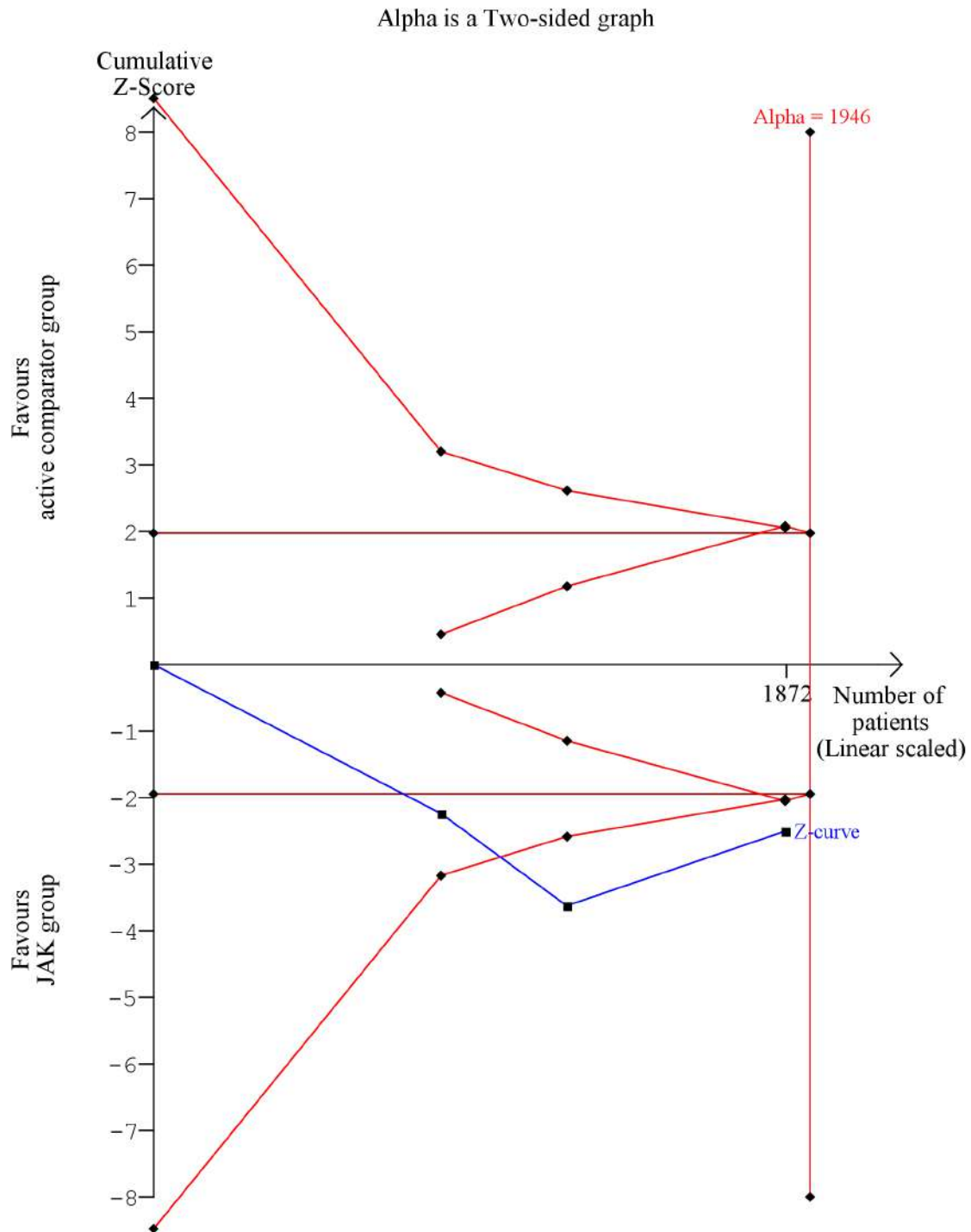
studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
FINCH 3 2019	1.90 (1.34, 2.70)	84/210	108/416	27.83
Conaghan 2016	2.46 (0.75, 8.11)	10/36	5/37	4.34
RA-BEGIN 2017	2.67 (1.69, 4.21)	67/159	45/210	20.64
SELECT-EARLY 2020	3.31 (2.31, 4.73)	221/631	44/314	27.37
ORAL Start 2014	3.48 (2.17, 5.57)	245/770	22/186	19.82
Overall (I-squared = 35.8%, p = 0.183)	2.71 (2.09, 3.50)	627/1806	224/1163	100.00

NOTE: Weights are from random effects analysis

favours active comparator group | favours JAK group
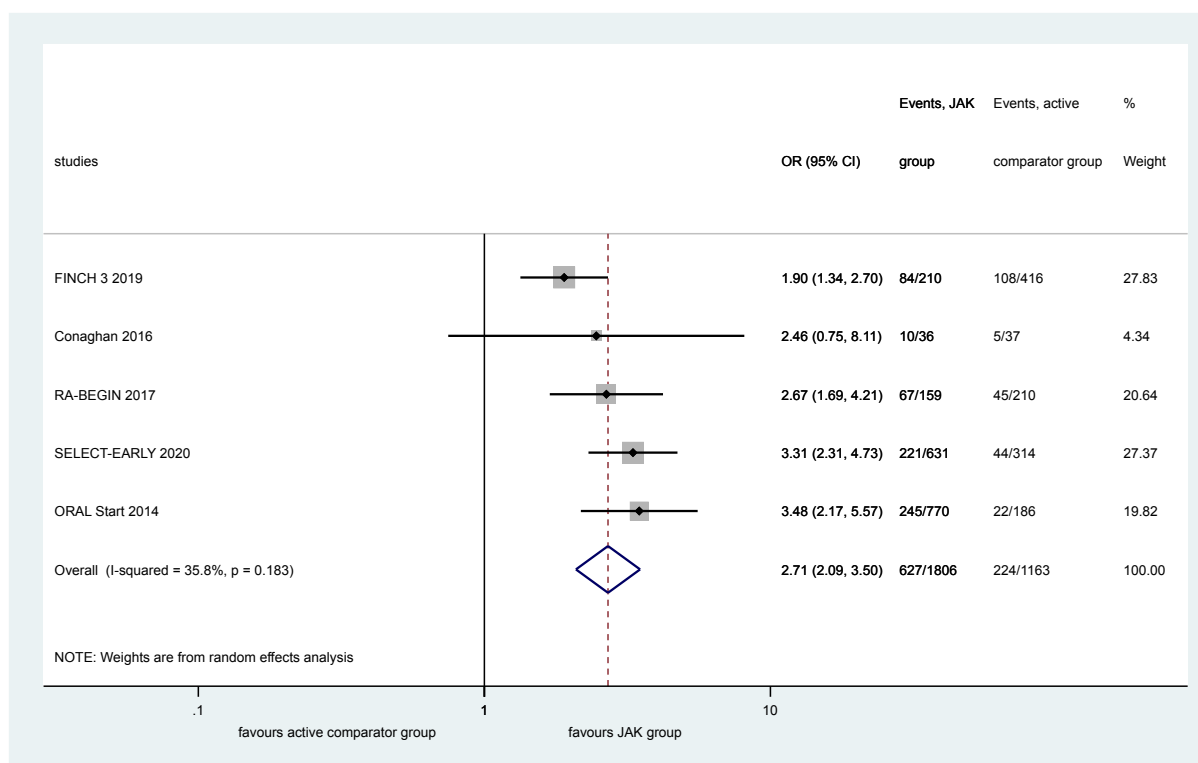
S135 b. Forest plot of studies comparing the number of patients reaching at least 70% in ACR criteria between patients treated with JAK inhibitors and MTX within 6 months (ACR70 response) - Secondary analysis only involving studies conducted in a MTX naïve population

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
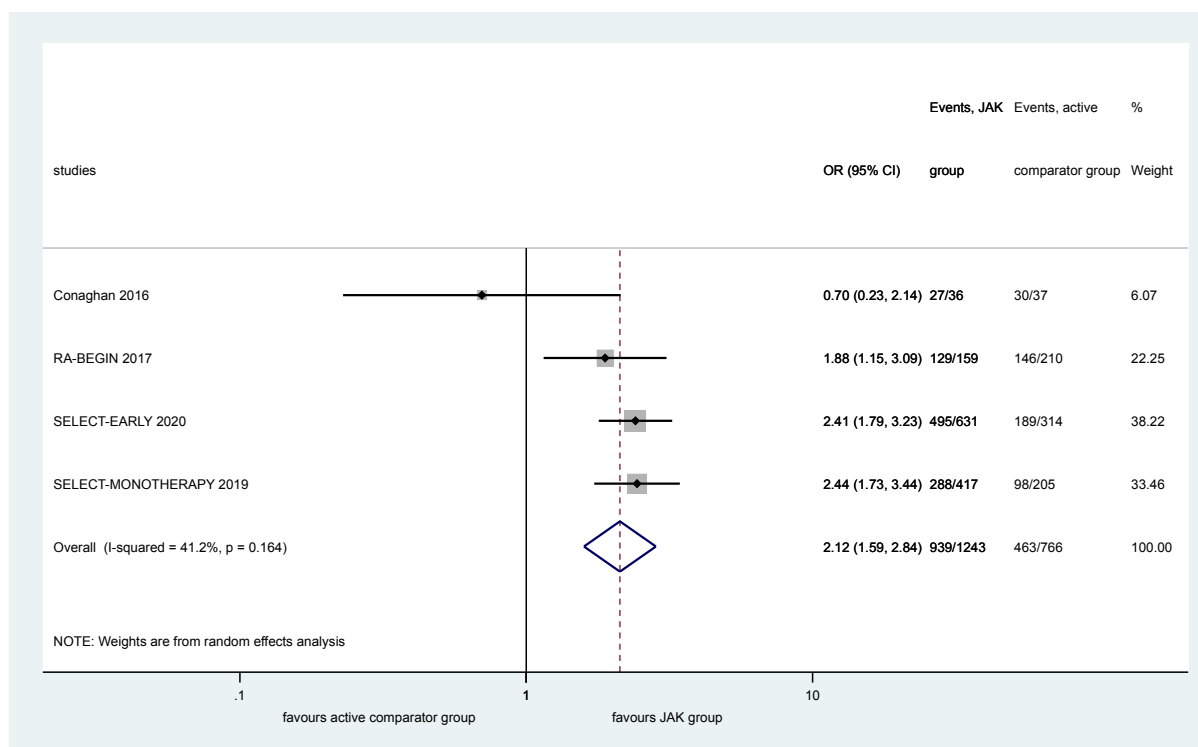
S136. Forest plot of studies comparing the number of patients reaching at least the minimum clinically important difference (≥0.22 MCID) in HAQ-DI between patients treated with JAK inhibitors and placebo within 6 months (HAQ-DI improvement)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S137. Forest plot of studies comparing the number of patients reaching remission defined by CDAI (≤ 2.8) between patients treated with JAK inhibitors and MTX within 6 months (CDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
RA-BEGIN 2017	1.37 (0.73, 2.58)	22/159	22/210	36.74
ORAL Start 2014	2.07 (1.18, 3.65)	126/698	15/156	37.51
SELECT-MONOTHERAPY 2019	20.34 (4.94, 83.80)	69/432	2/216	25.75
Overall (I-squared = 85.4%, p = 0.001)	3.21 (0.99, 10.40)	217/1289	39/582	100.00

NOTE: Weights are from random effects analysis

.1 — 1 — 10 — 100
favours active comparator group — favours JAK group
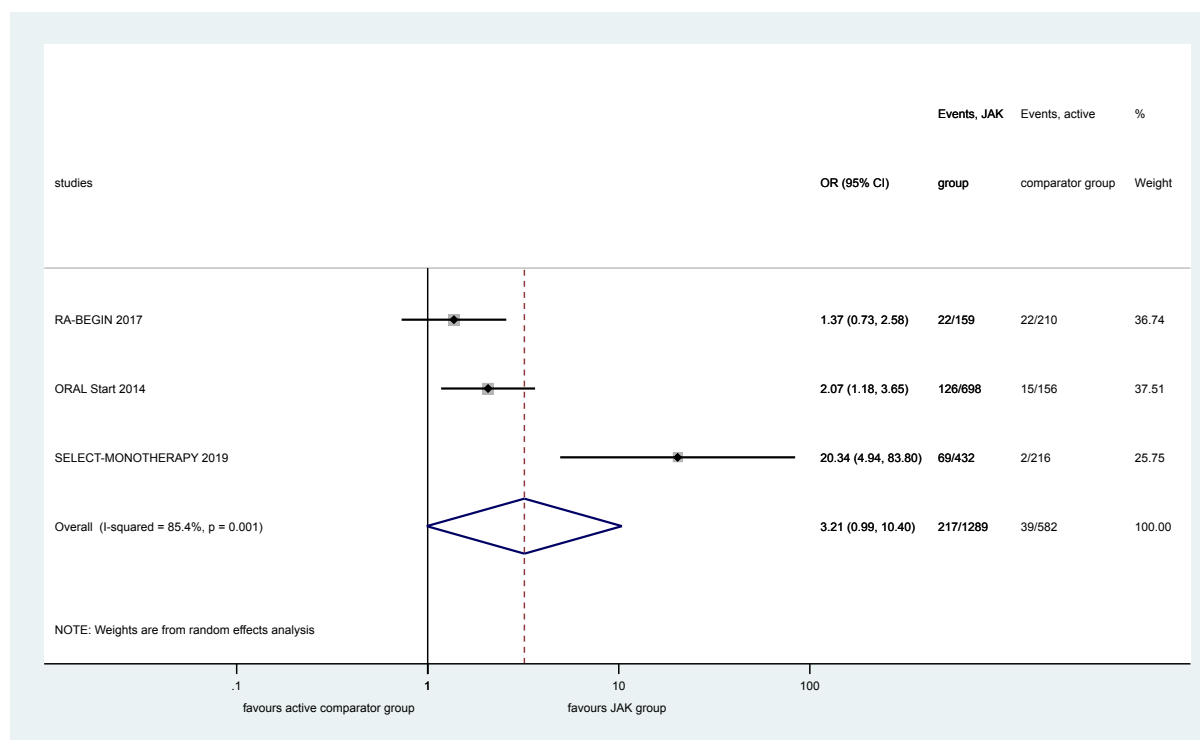
S138.Forest plot of studies comparing the number of patients reaching remission defined by SDAI (≤ 3.3) between patients treated with JAK inhibitors and MTX within 6 months (SDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S7139 Forest plot of studies comparing the number of patients reaching remission defined by DAS28-CRP (<2.6) between patients treated with JAK inhibitors and MTX within 6 months (DAS28-CRP remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
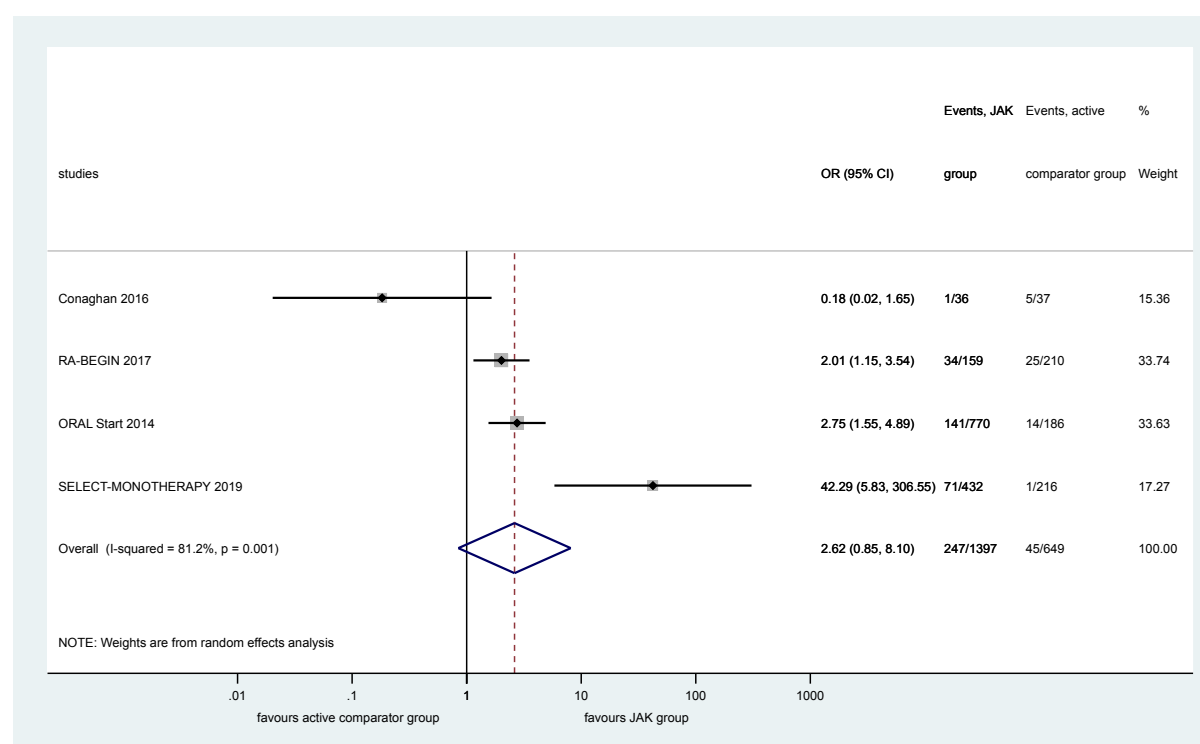
S140 a. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-ESR (<2.6) between patients treated with JAK inhibitors and MTX within 6 months (DAS28-ESR remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
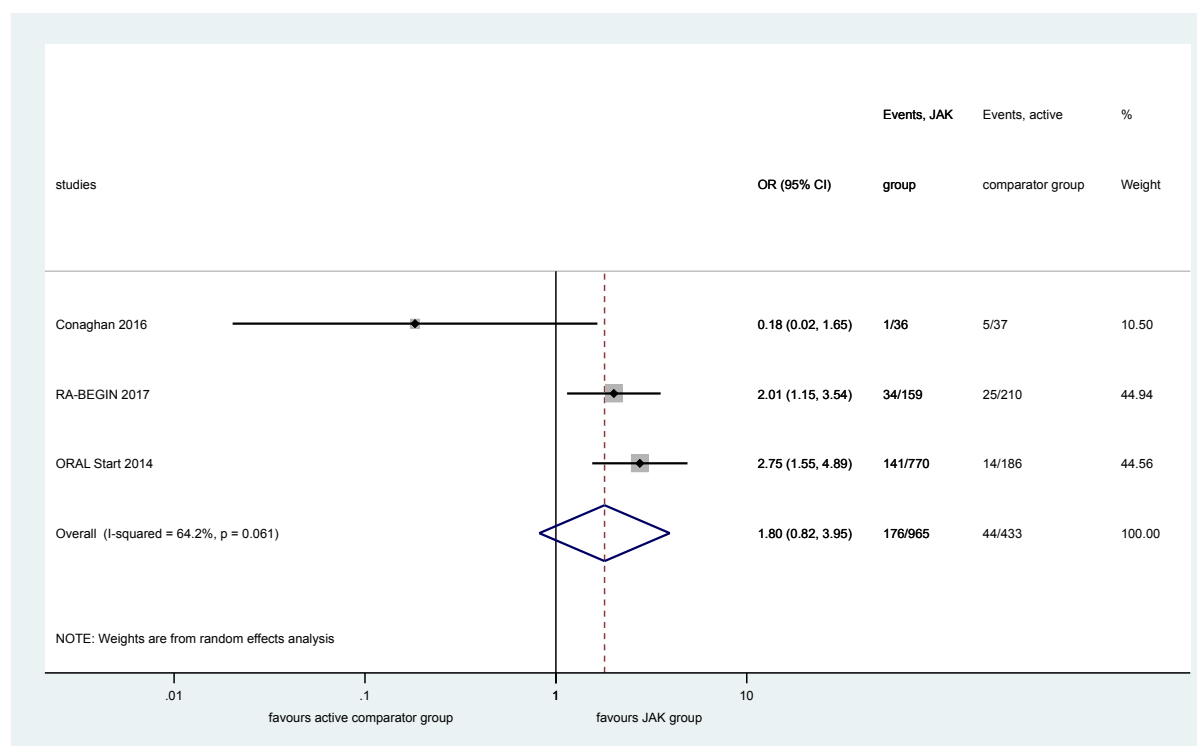
S140 b. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-ESR (<2.6) between patients treated with JAK inhibitors and MTX within 6 months (DAS28-ESR remission) - Secondary analysis only involving studies conducted in a MTX naïve population

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
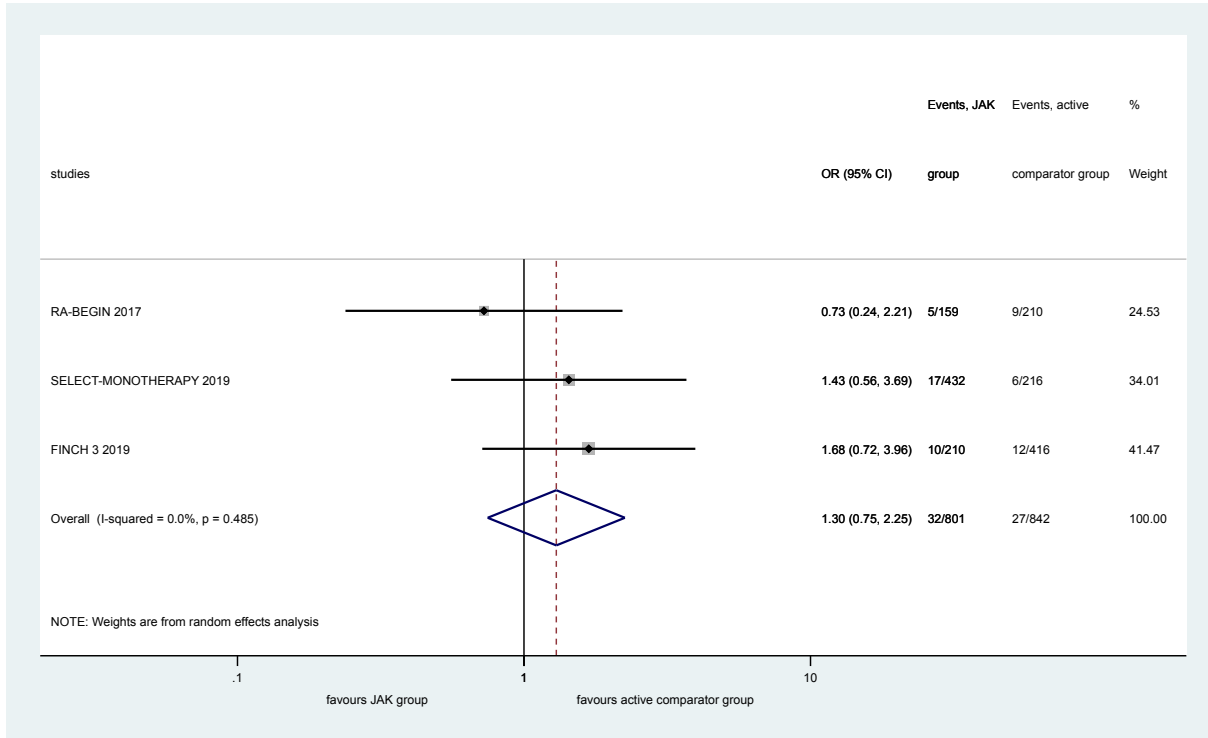
S141. Forest plot of studies comparing the number of patients experiencing serious side effects during the study between patients treated with JAK inhibitors and MTX within 6 months (Serious side effects)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S142. Forest plot of studies comparing the LSM change from baseline in the score on HAQ-DI between patients treated with JAK inhibitors and MTX within 6 months (HAQ-DI difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
FINCH 3 2019	0.10 (-1.92, 2.12)	181, 10.2 (11.4)	365, 10.1 (11.2)	30.97
ORAL Start 2014	4.76 (3.43, 6.09)	731, 9.82 (8.14)	169, 5.06 (7.93)	34.54
SELECT-MONOTHERAPY 2019	4.95 (3.61, 6.29)	432, 9.25 (8.25)	216, 4.3 (8.2)	34.49
Overall (I-squared = 88.6%, p = 0.000)	3.38 (0.76, 6.00)	1344	750	100.00

NOTE: Weights are from random effects analysis

-2   0   2   8

favours active comparator group   favours JAK group

S143. Forest plot of studies comparing the LSM change from baseline in the score on SF-36 assessing the physical component score between patients treated with JAK inhibitors and MTX within 6 months (SF-36 PCS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
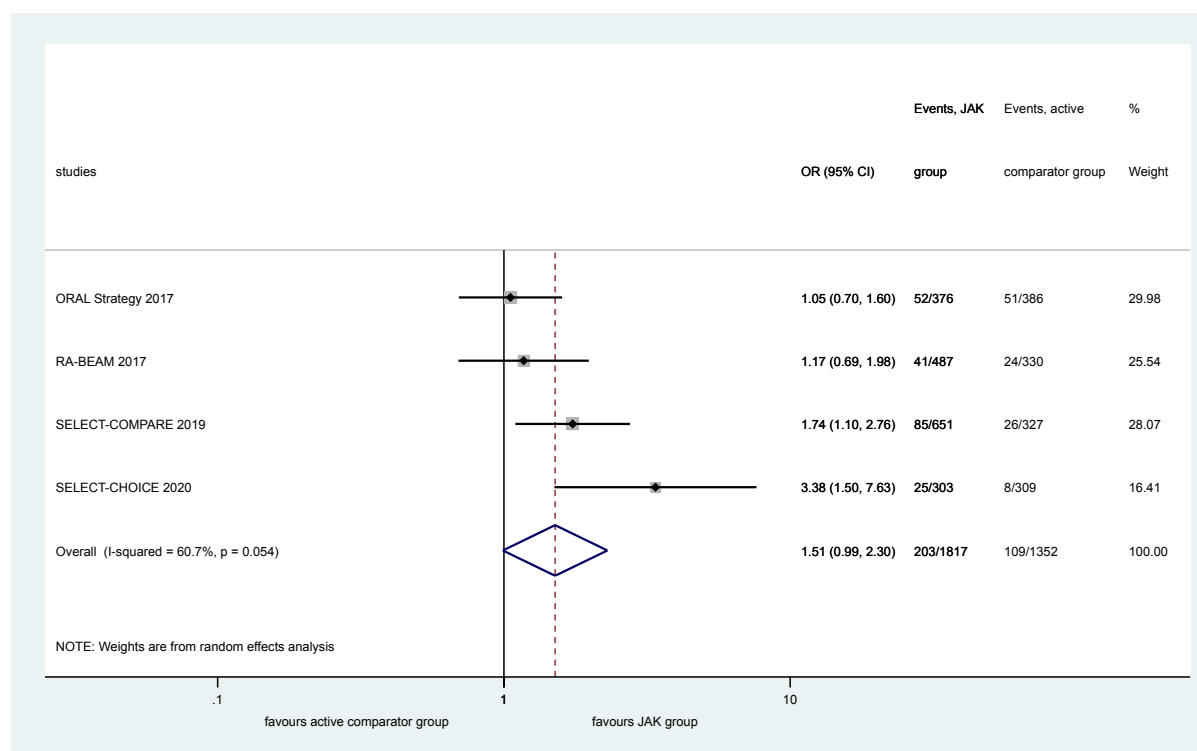
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
FINCH 3 2019	0.10 (-1.92, 2.12)	181, 10.2 (11.4)	365, 10.1 (11.2)	21.91
SELECT-EARLY 2020	2.99 (1.66, 4.32)	598, 9.79 (9.37)	277, 6.8 (9.34)	28.23
ORAL Start 2014	3.14 (1.67, 4.61)	731, 8.47 (8.98)	169, 5.33 (8.71)	26.96
RA-BEGIN 2017	4.40 (2.50, 6.30)	159, 13.3 (9.26)	210, 8.9 (9.19)	22.91
Overall (I-squared = 69.7%, p = 0.019)	2.72 (1.22, 4.22)	1669	1021	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S144. Forest plot of studies comparing the LSM change from baseline in the score on FACIT-F between patients treated with JAK inhibitors and MTX within 6 months (FACIT-F)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
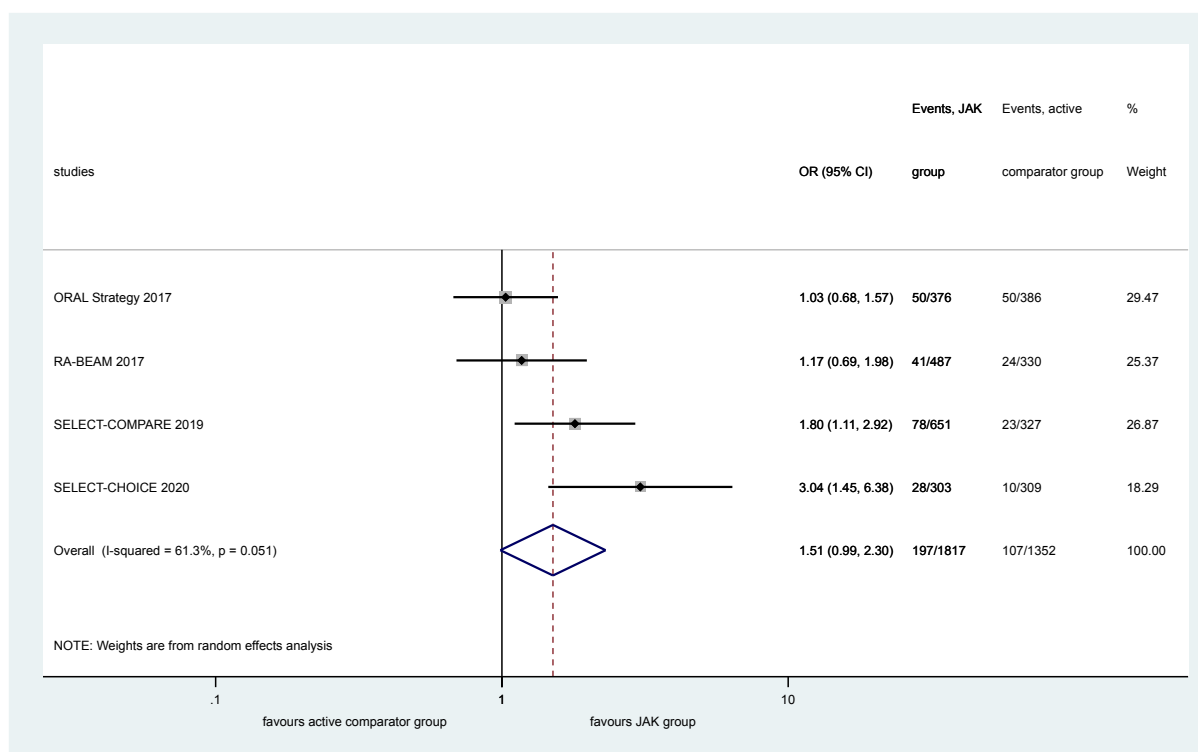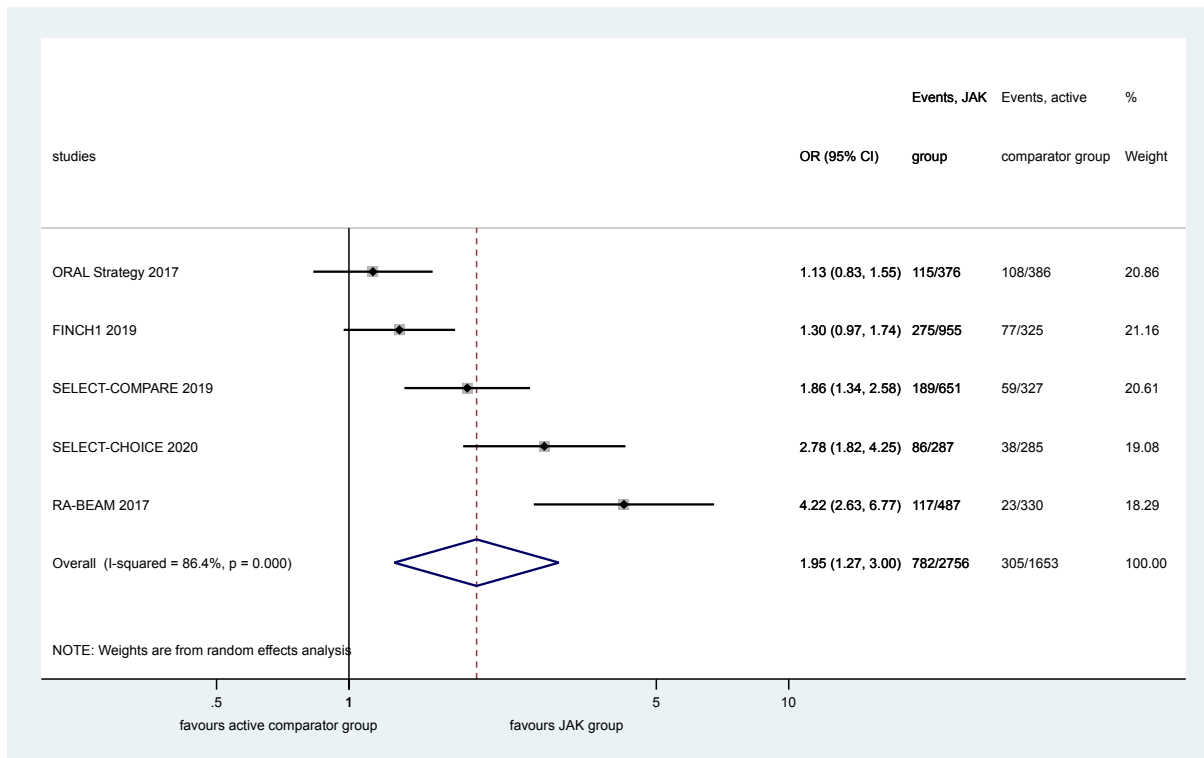
Alpha is a Two-sided graph

S145. TSA of ACR70 response (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.

S146. TSA of HAQ-DI improvement (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.

S147. TSA of CDAI remission (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards) but does not reach the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome but the addition of further studies might influence these results.

Alpha is a Two-sided graph
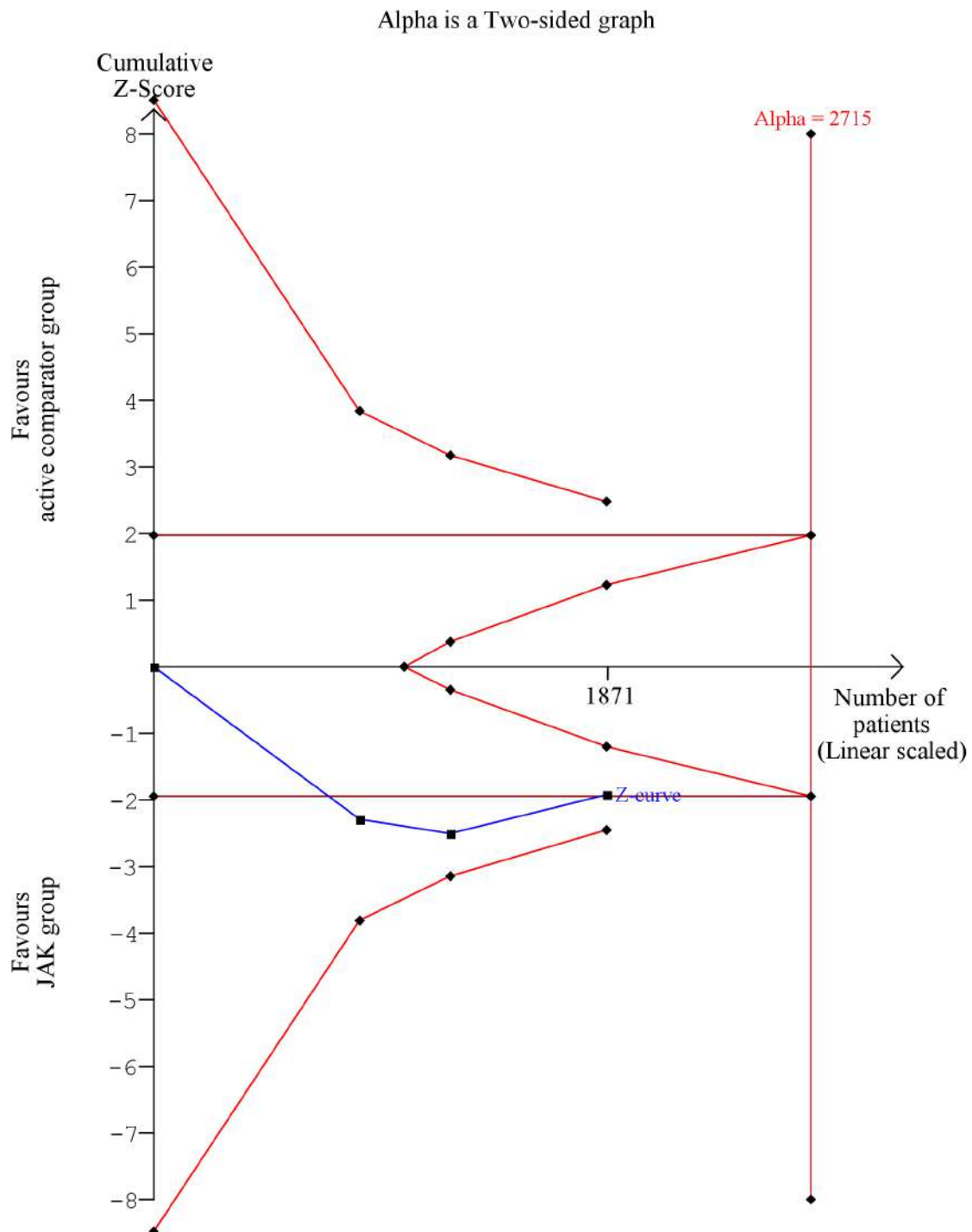
S148. TSA of SDAI remission (JAK-inhibitors vs MTX)

The blue line (Z-curve) first crosses the conventional boundary (brown straight line) but returns and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome but the addition of further studies might influence these results.

S149. TSA of DAS28-CRP remission (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.
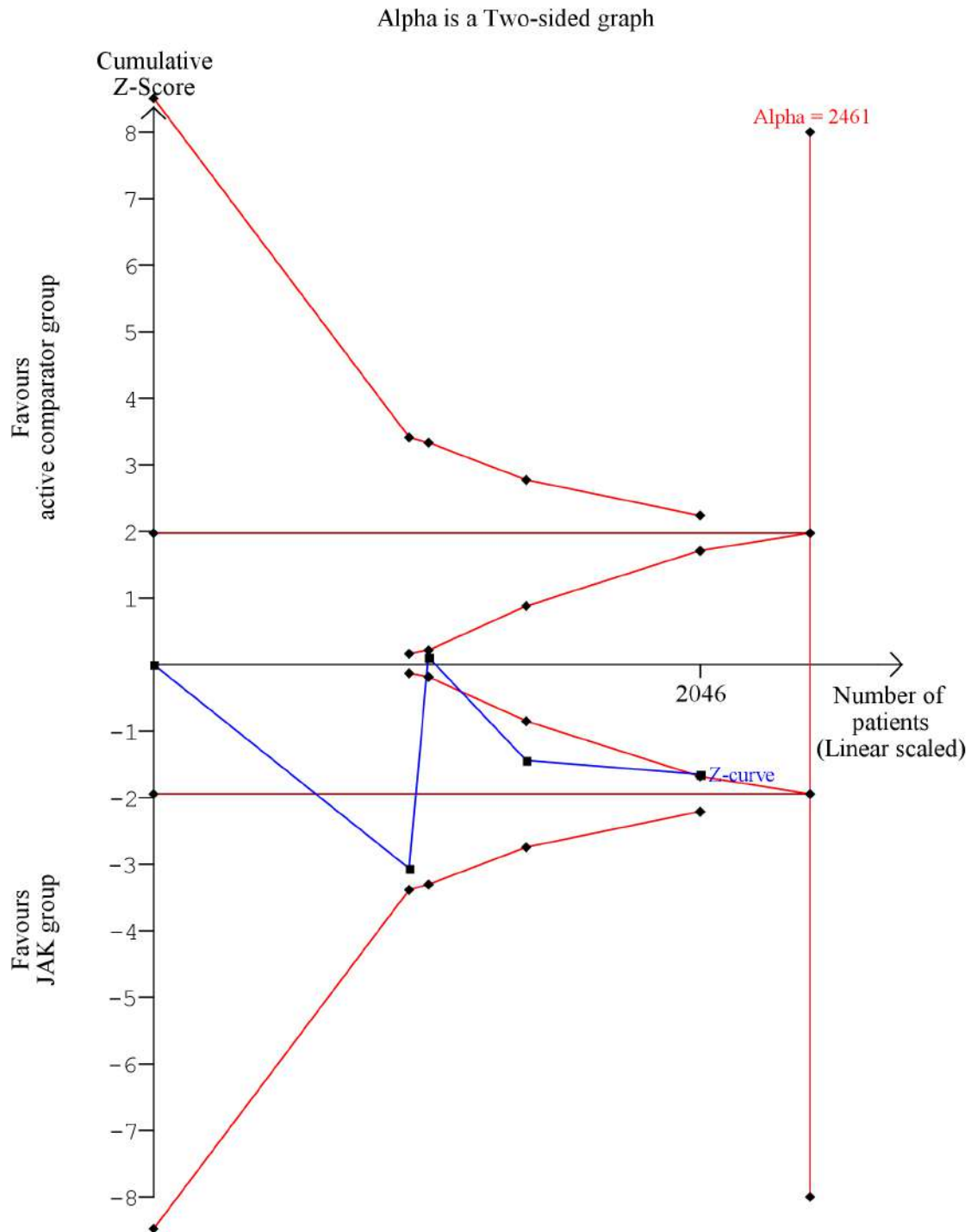
S150. TSA of DAS28-ESR remission (JAK-inhibitors vs MTX)

The blue line (Z-curve) first crosses the conventional boundary (brown straight line) but returns and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome but the addition of further studies might influence these results.

Alpha is a Two-sided graph
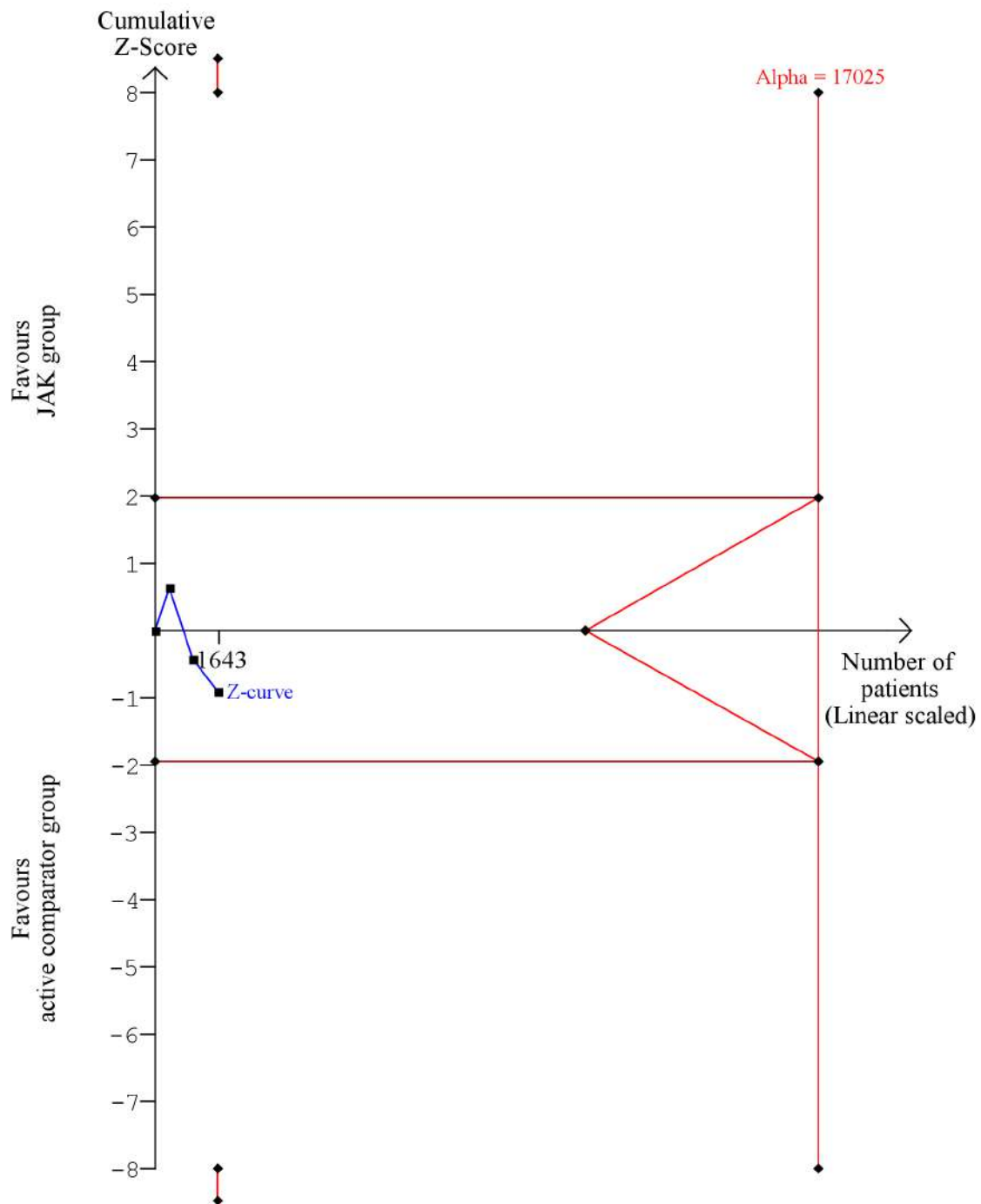
S151. TSA of serious side effects (JAK-inhibitors vs MTX)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results
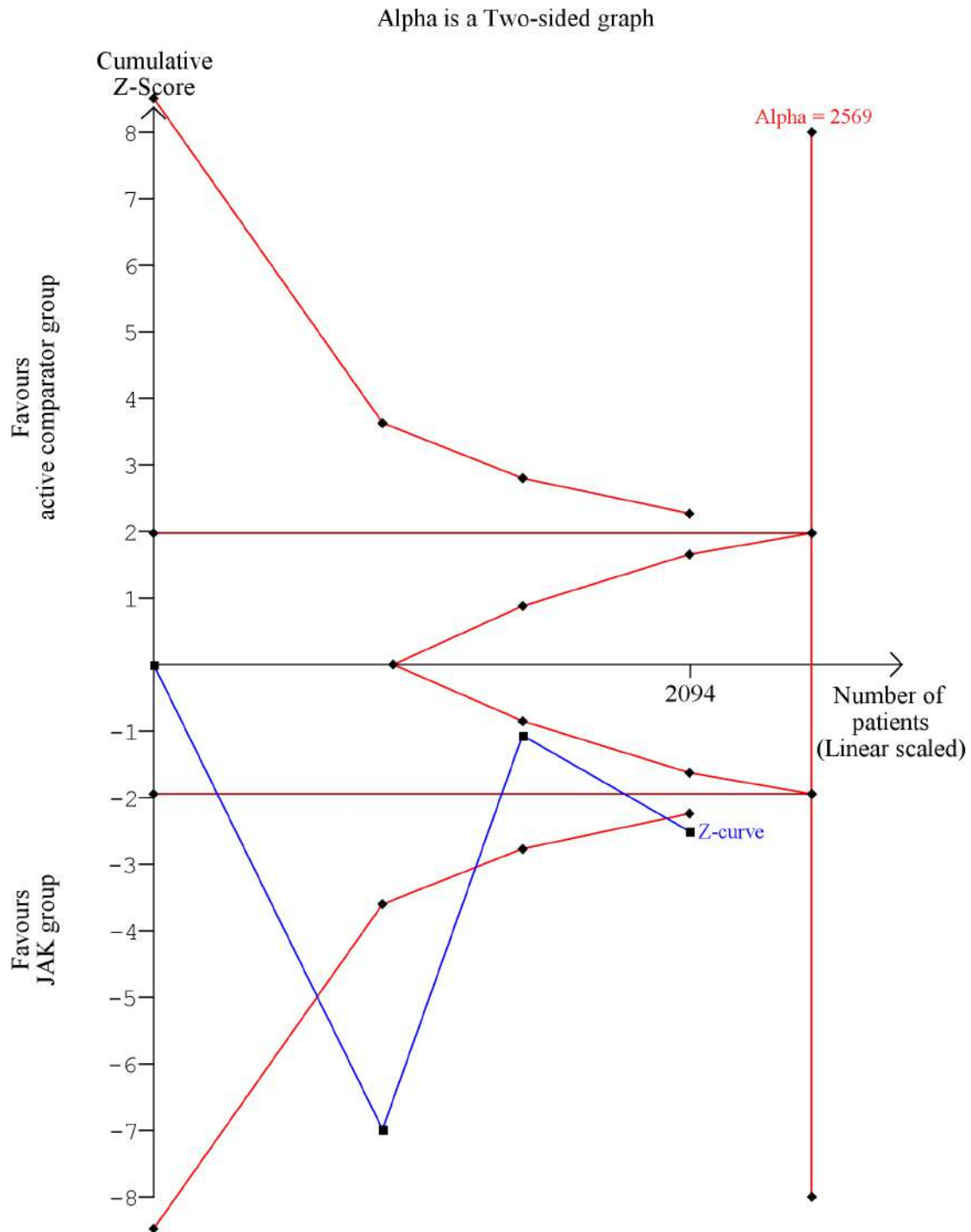
S152. TSA of SF-36 PCS (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards) but does not reach the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome but the addition of further studies might influence these results.
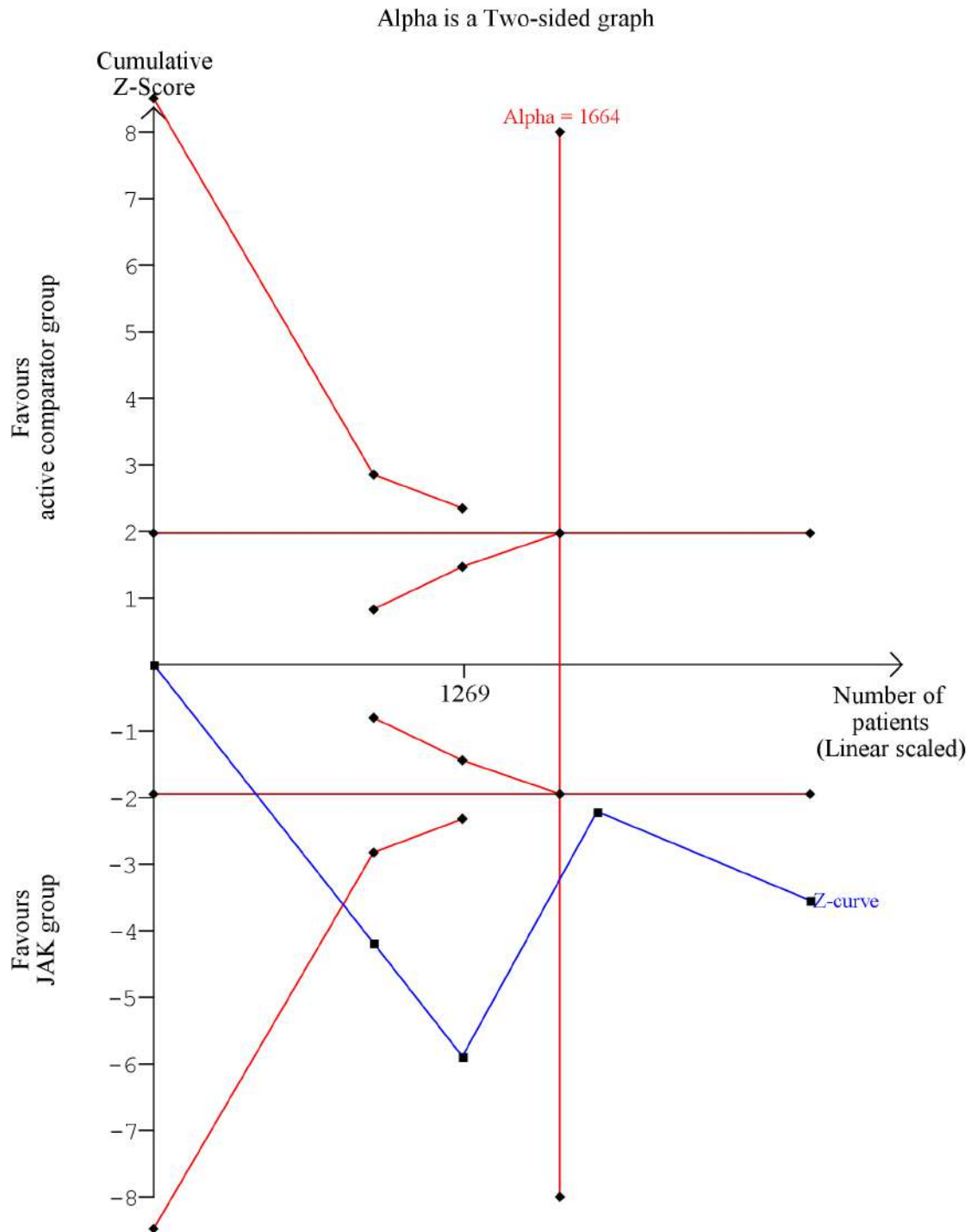
S153. TSA of FACIT-F (JAK-inhibitors vs MTX)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

(3) JAK inhibitors compared to bDMARDs

studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
ORAL Strategy 2017	1.11 (0.81, 1.53)	275/376	274/386	15.65
FINCH1 2019	1.13 (0.85, 1.49)	699/955	230/325	20.22
ORAL Standard 2012	1.21 (0.86, 1.71)	204/392	94/199	13.45
SELECT-COMPARE 2019	1.44 (1.08, 1.90)	462/651	206/327	19.86
RA-BEAM 2017	1.45 (1.08, 1.95)	339/487	202/330	18.19
SELECT-CHOICE 2020	1.57 (1.10, 2.23)	229/303	205/309	12.62
Overall (I-squared = 0.0%, p = 0.520)	1.30 (1.15, 1.48)	2208/3164	1211/1876	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S154. Forest plot of studies comparing the number of patients reaching at least 20% in ACR criteria between patients treated with JAK inhibitors and bDMARDs within 6 months (ACR20 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
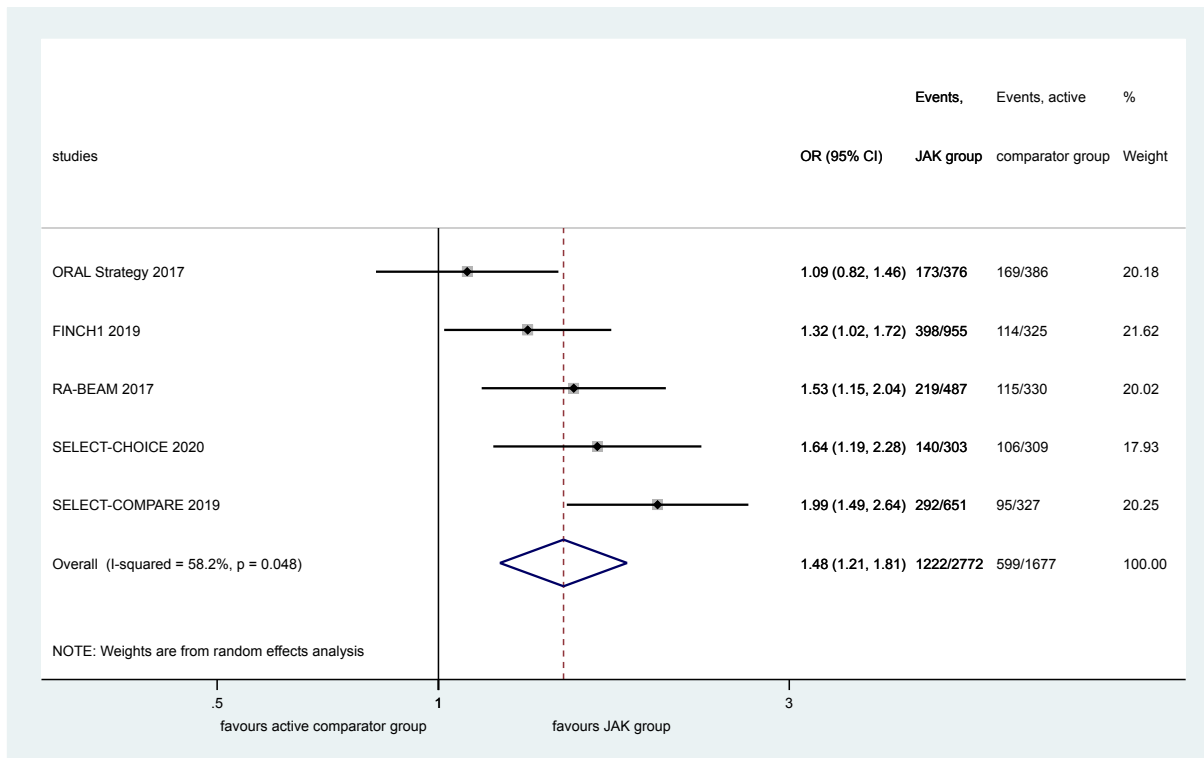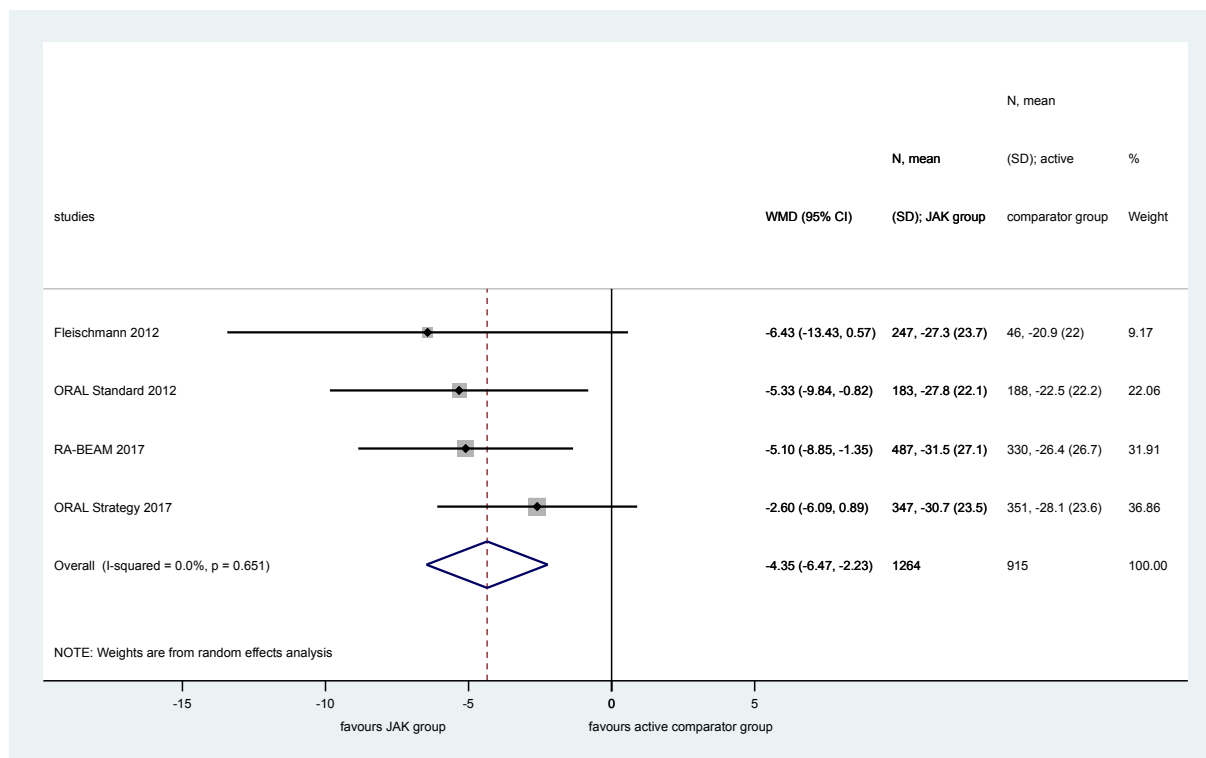
S155. Forest plot of studies comparing the number of patients reaching at least 50% in ACR criteria between patients treated with JAK inhibitors and bDMARDs within 6 months (ACR50 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
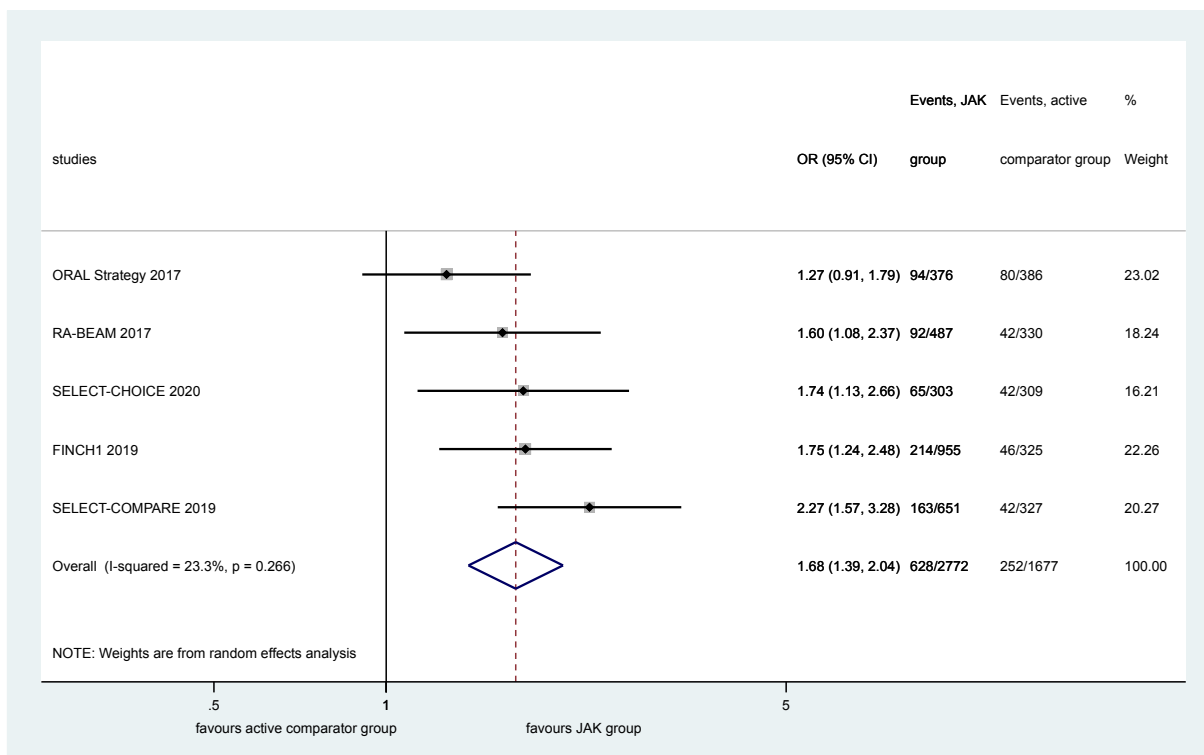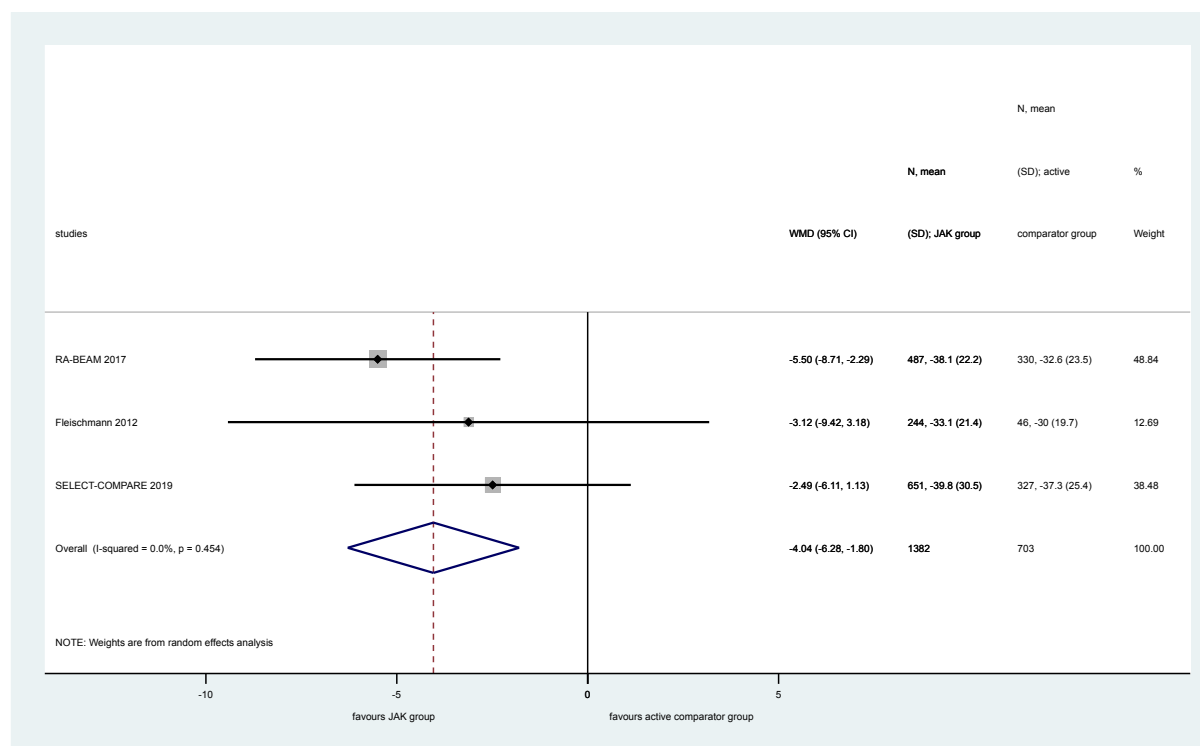
S156. Forest plot of studies comparing the number of patients reaching at least 70% in ACR criteria between patients treated with JAK inhibitors and bDMARDs within 6 months (ACR70 response)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
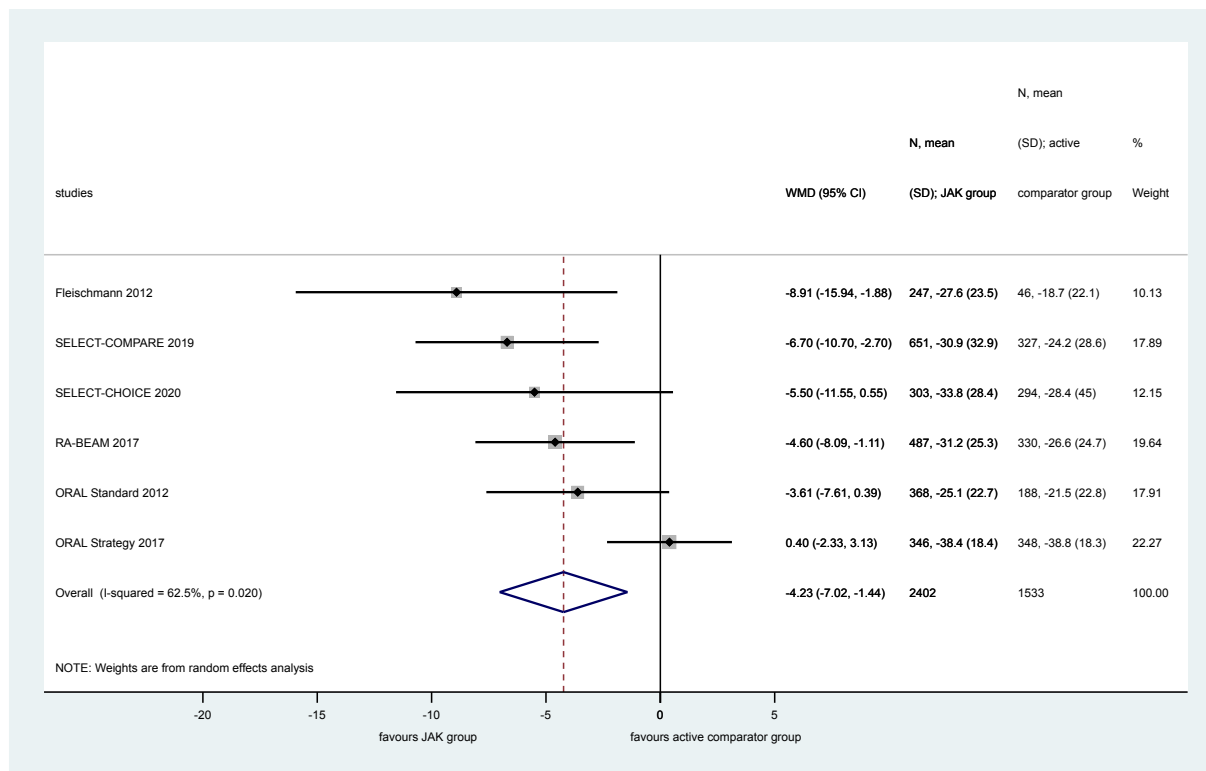

S157. Forest plot of studies comparing the number of patients reaching remission defined by CDAI (≤ 2.8) between patients treated with JAK inhibitors and bDMARDs within 6 months (CDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI) and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
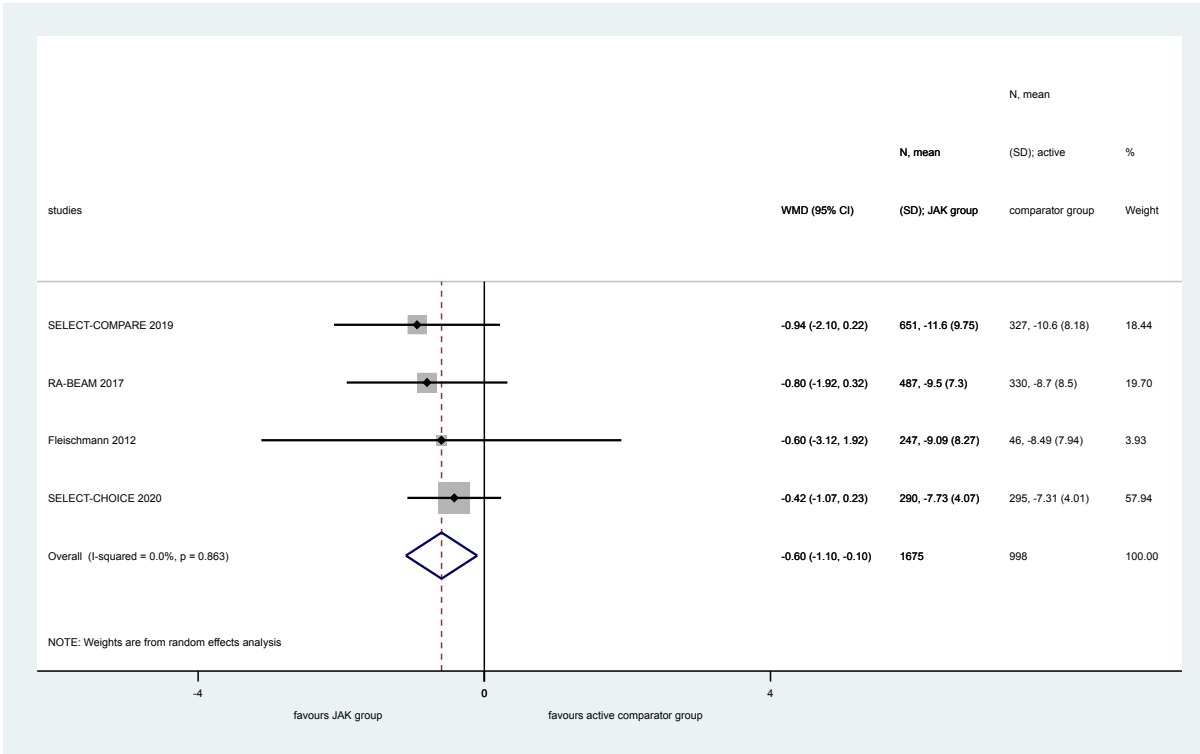
studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
ORAL Strategy 2017	1.03 (0.68, 1.57)	50/376	50/386	29.47
RA-BEAM 2017	1.17 (0.69, 1.98)	41/487	24/330	25.37
SELECT-COMPARE 2019	1.80 (1.11, 2.92)	78/651	23/327	26.87
SELECT-CHOICE 2020	3.04 (1.45, 6.38)	28/303	10/309	18.29
Overall (I-squared = 61.3%, p = 0.051)	1.51 (0.99, 2.30)	197/1817	107/1352	100.00

NOTE: Weights are from random effects analysis

favours active comparator group favours JAK group

S158. Forest plot of studies comparing the number of patients reaching remission defined by SDAI (≤ 3.3) between patients treated with JAK inhibitors and bDMARDs within 6 months (SDAI remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
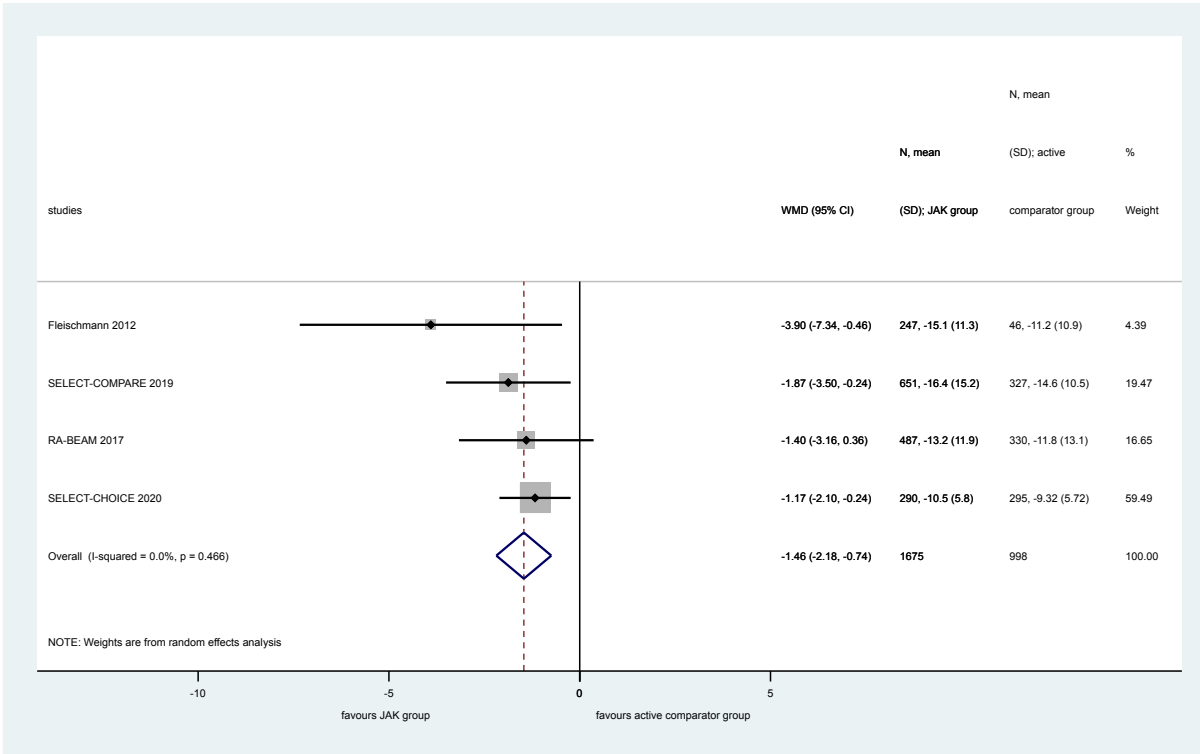
S159. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-CRP (<2.6) between patients treated with JAK inhibitors and tbDMRADs within 6 months (DAS28-CRP remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
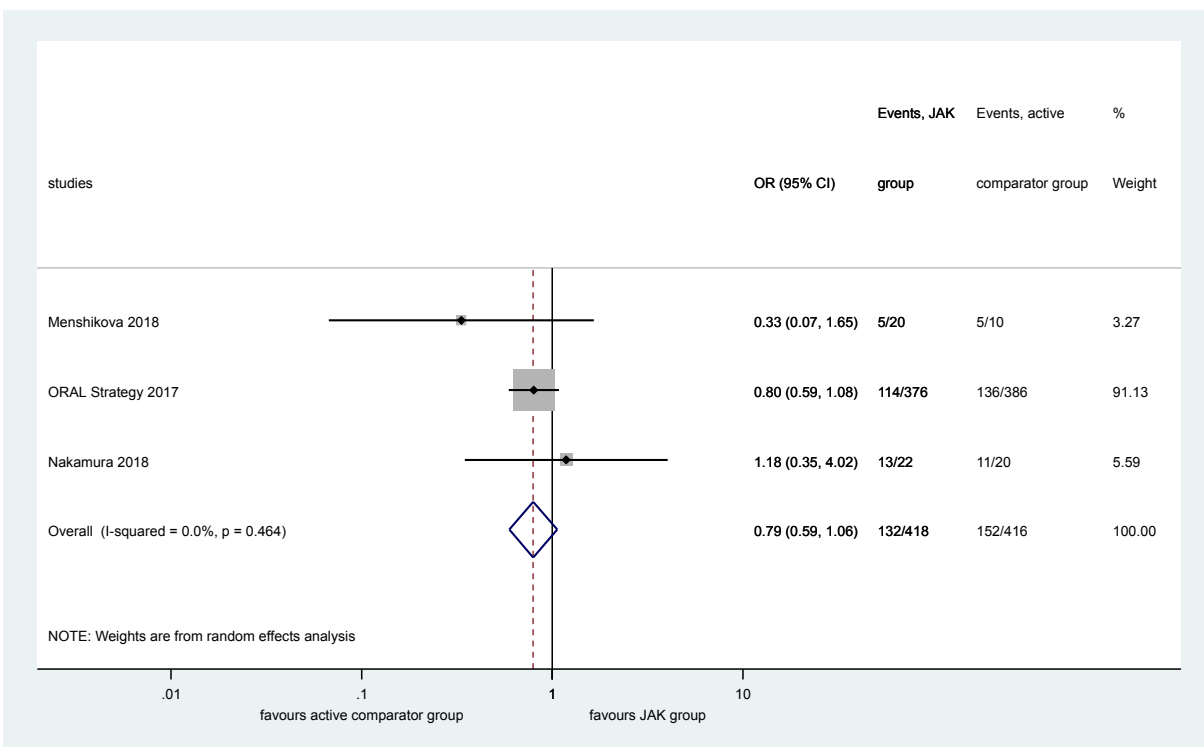
S160. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-CRP (<2.6) between patients treated with JAK inhibitors and tbDMRADs over 6 months (DAS28-CRP remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.


studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
RA-BEAM 2017	0.90 (0.58, 1.40)	51/487	38/330	26.92
ORAL Strategy 2017	0.96 (0.62, 1.48)	45/376	48/386	27.15
ORAL Standard 2012	1.31 (0.67, 2.55)	33/353	13/178	22.34
SELECT-CHOICE 2020	3.42 (1.86, 6.28)	45/303	15/309	23.59
Overall (I-squared = 78.8%, p = 0.003)	1.36 (0.77, 2.41)	174/1519	114/1203	100.00

NOTE: Weights are from random effects analysis

.1   1   10

favours active comparator group   favours JAK group

S161. Forest plot of studies comparing the number of patients reaching remission defined by DAS28-ESR (<2.6) between patients treated with JAK inhibitors and tbDMRADs within 6 months (DAS28-ESR remission)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
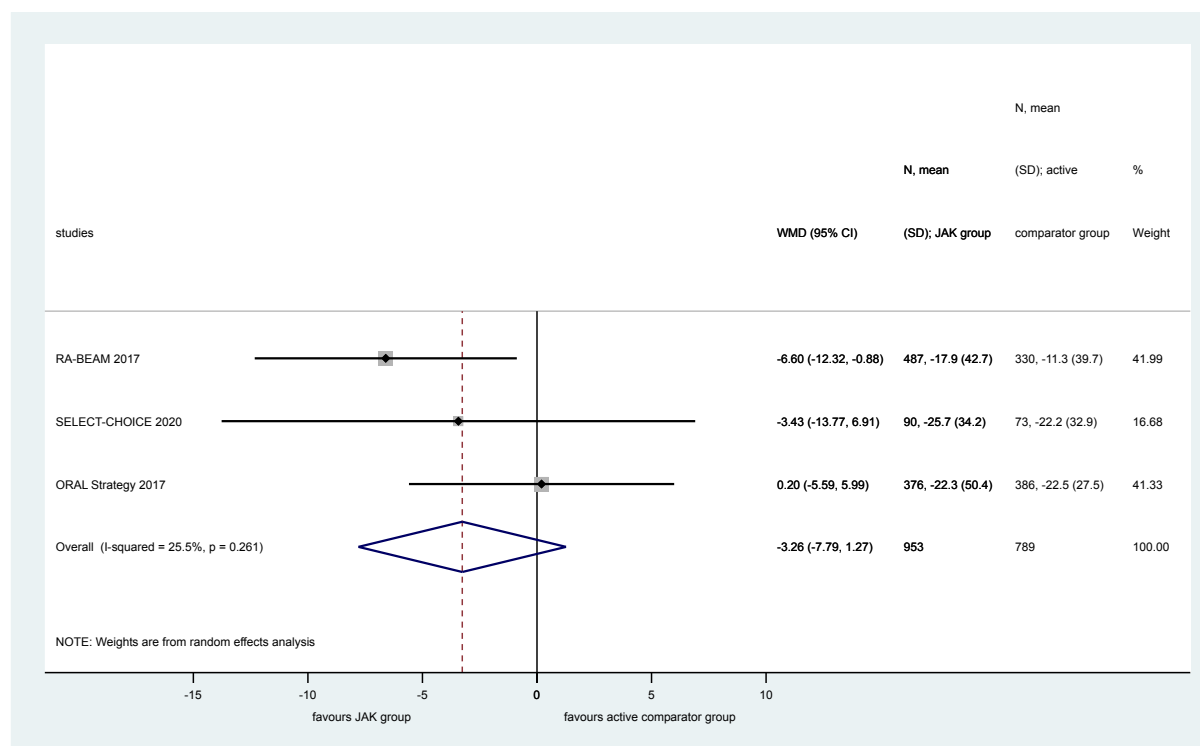
studies	OR (95% CI)	Events, JAK group	Events, active comparator group	% Weight
FINCH1 2019	1.10 (0.59, 2.03)	45/955	14/325	58.84
ORAL Standard 2012	2.29 (0.85, 6.13)	22/405	5/204	31.90
Fleischmann 2012	4.00 (0.52, 30.74)	17/238	1/53	9.25
Overall (I-squared = 23.3%, p = 0.272)	1.56 (0.82, 2.98)	84/1598	20/582	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours active comparator group

S162. Forest plot of studies comparing the number of patients experiencing serious side effects during the study between patients treated with JAK inhibitors and bDMARDs within 6 months (Serious side effects)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
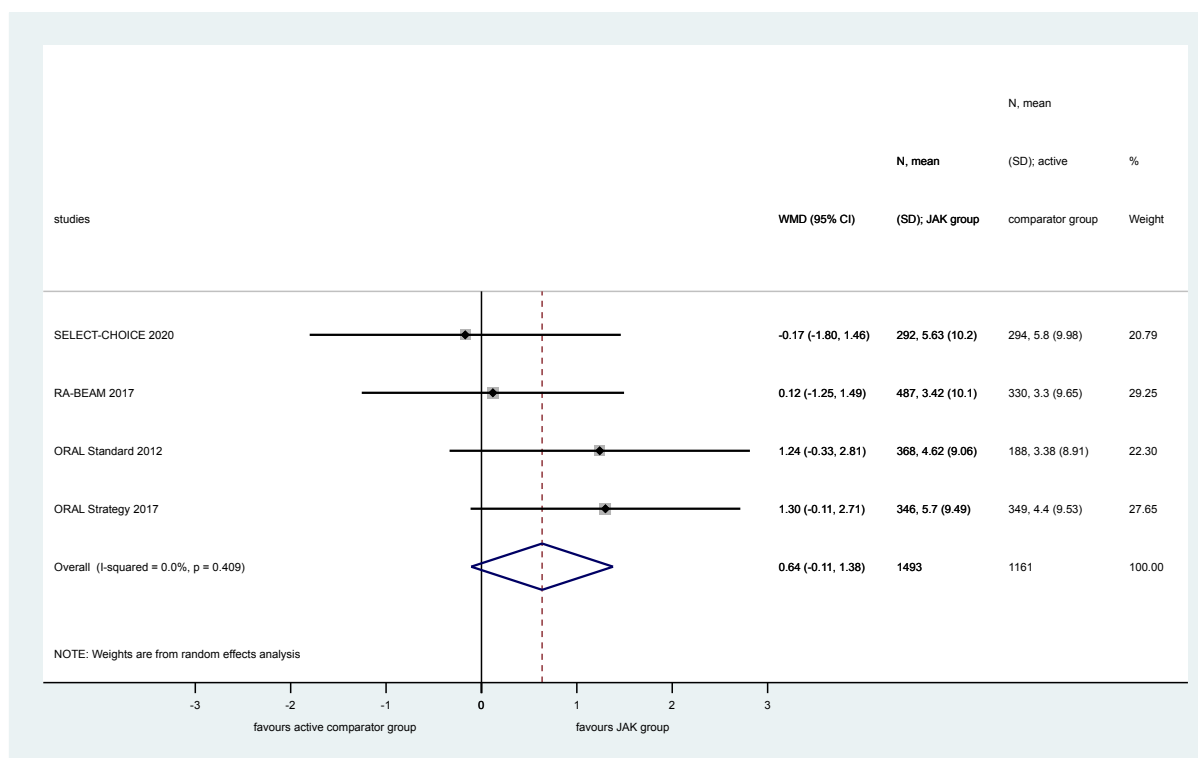
S163. Forest plot of studies comparing the number of patients who dies during study period between patients treated with JAK inhibitors and bDMARDs (Deaths)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S164. Forest plot of studies comparing the LSM change from baseline in pain measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and bDMARDs within 6 months (Pain VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
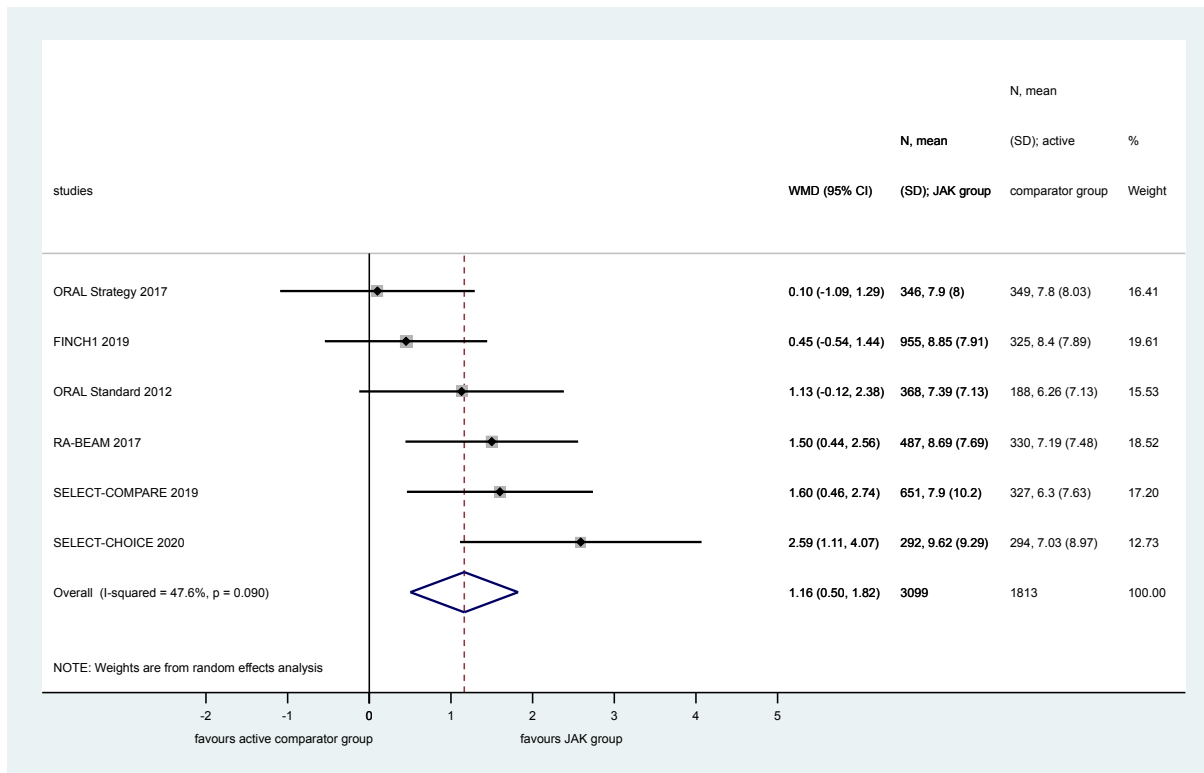

S165. Forest plot of studies comparing the LSM change from baseline in PGA measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and bDMARDs within 6 months (PGA VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
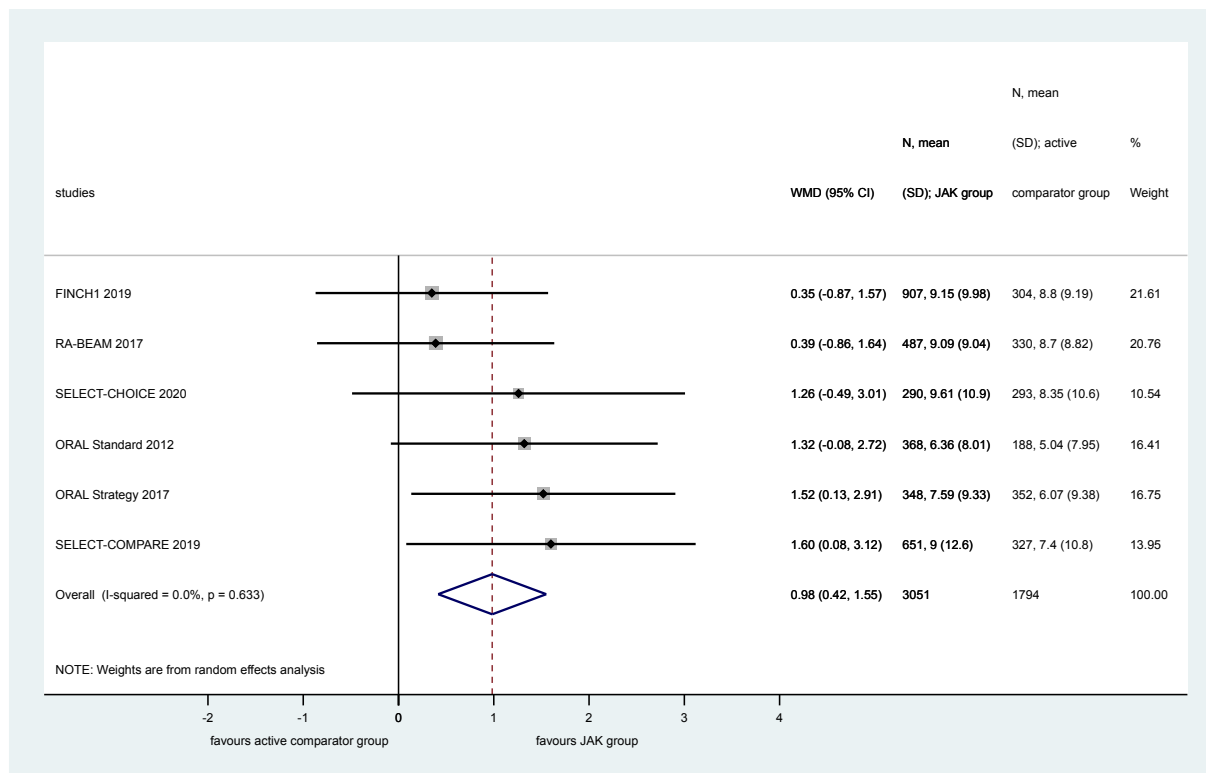
studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
Fleischmann 2012	-8.91 (-15.94, -1.88)	247, -27.6 (23.5)	46, -18.7 (22.1)	10.13
SELECT-COMPARE 2019	-6.70 (-10.70, -2.70)	651, -30.9 (32.9)	327, -24.2 (28.6)	17.89
SELECT-CHOICE 2020	-5.50 (-11.55, 0.55)	303, -33.8 (28.4)	294, -28.4 (45)	12.15
RA-BEAM 2017	-4.60 (-8.09, -1.11)	487, -31.2 (25.3)	330, -26.6 (24.7)	19.64
ORAL Standard 2012	-3.61 (-7.61, 0.39)	368, -25.1 (22.7)	188, -21.5 (22.8)	17.91
ORAL Strategy 2017	0.40 (-2.33, 3.13)	346, -38.4 (18.4)	348, -38.8 (18.3)	22.27
Overall (I-squared = 62.5%, p = 0.020)	-4.23 (-7.02, -1.44)	2402	1533	100.00

NOTE: Weights are from random effects analysis

favours JAK group   favours active comparator group

S166. Forest plot of studies comparing the LSM change from baseline in PtGA measured on VAS ranging from 0-100 mm between patients treated with JAK inhibitors and bDMARDs within 6 months (PtGA VAS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
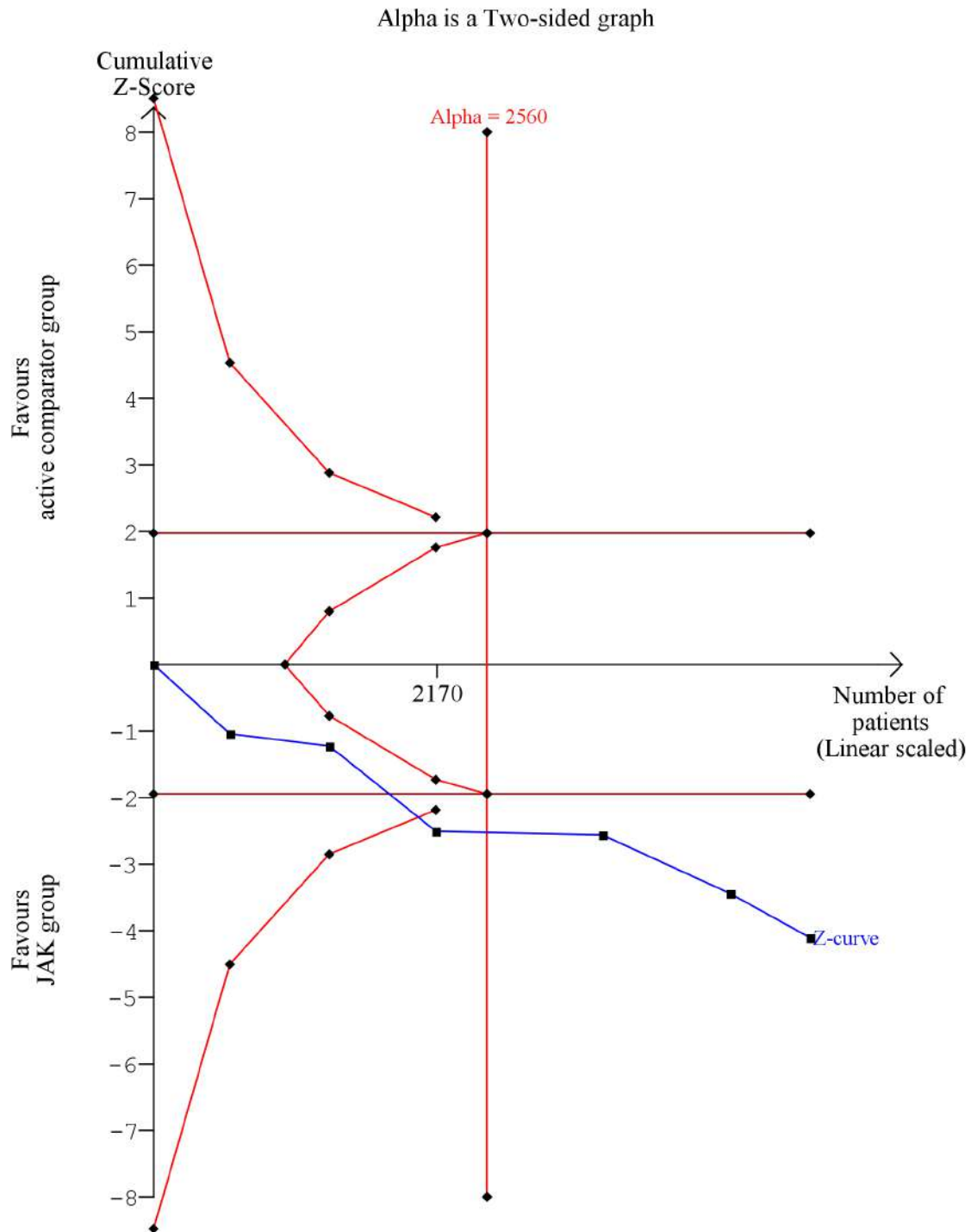
S167. Forest plot of studies comparing the LSM change from baseline in the number of swollen joint counts between patients treated with JAK inhibitors and bDMARDs within 6 months (Swollen Joint Count)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
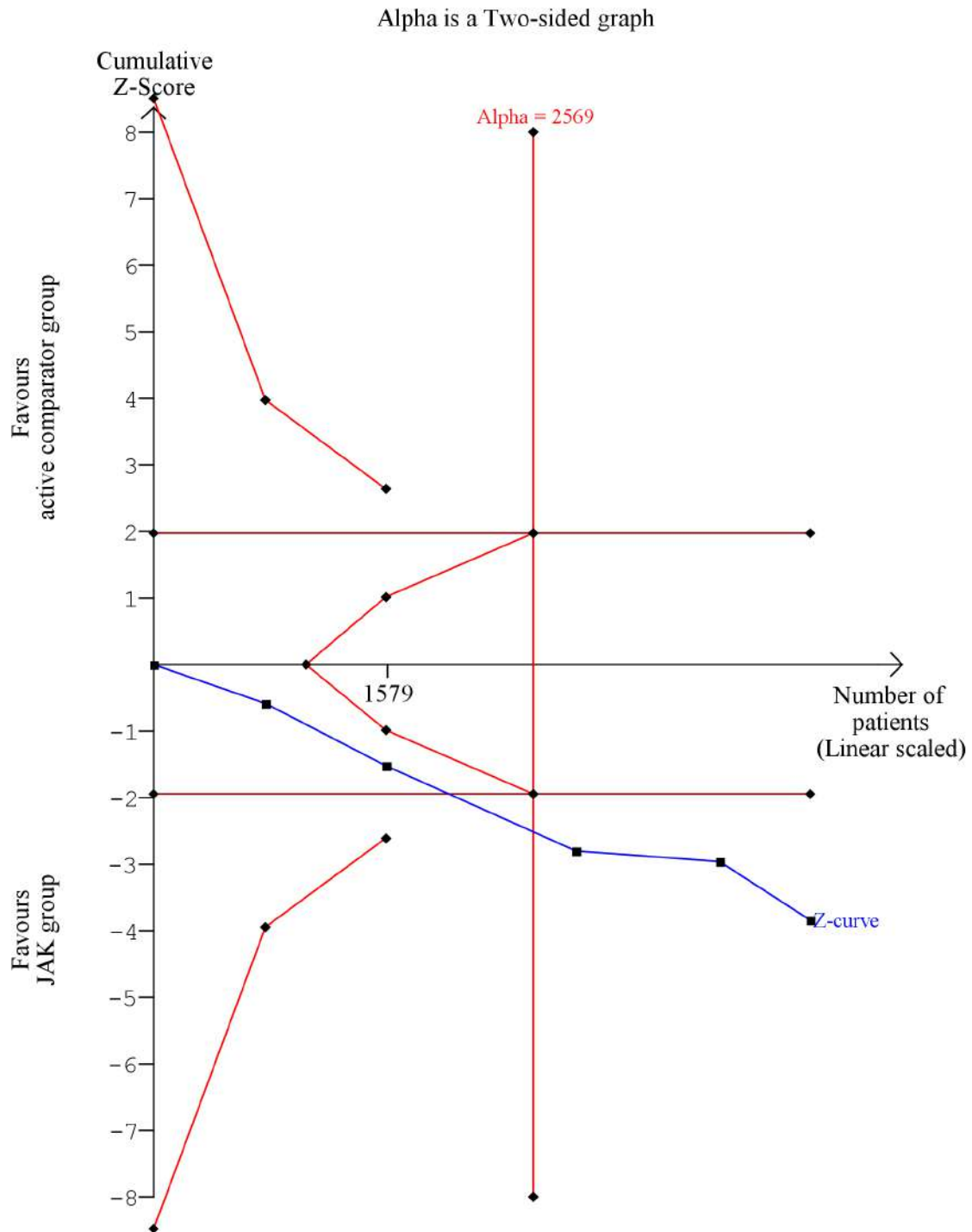
S168. Forest plot of studies comparing the LSM change from baseline in the number of tender joint counts between patients treated with JAK inhibitors and bDMARDs within 6 months (Tender Joint Count)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
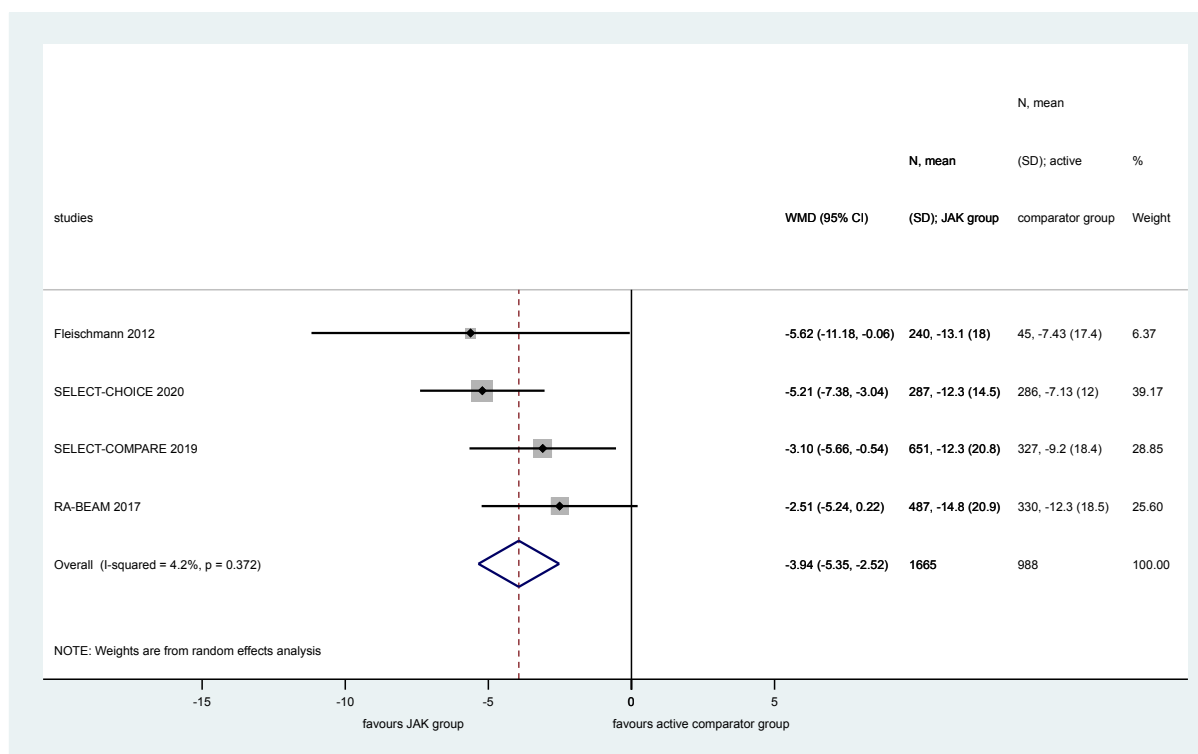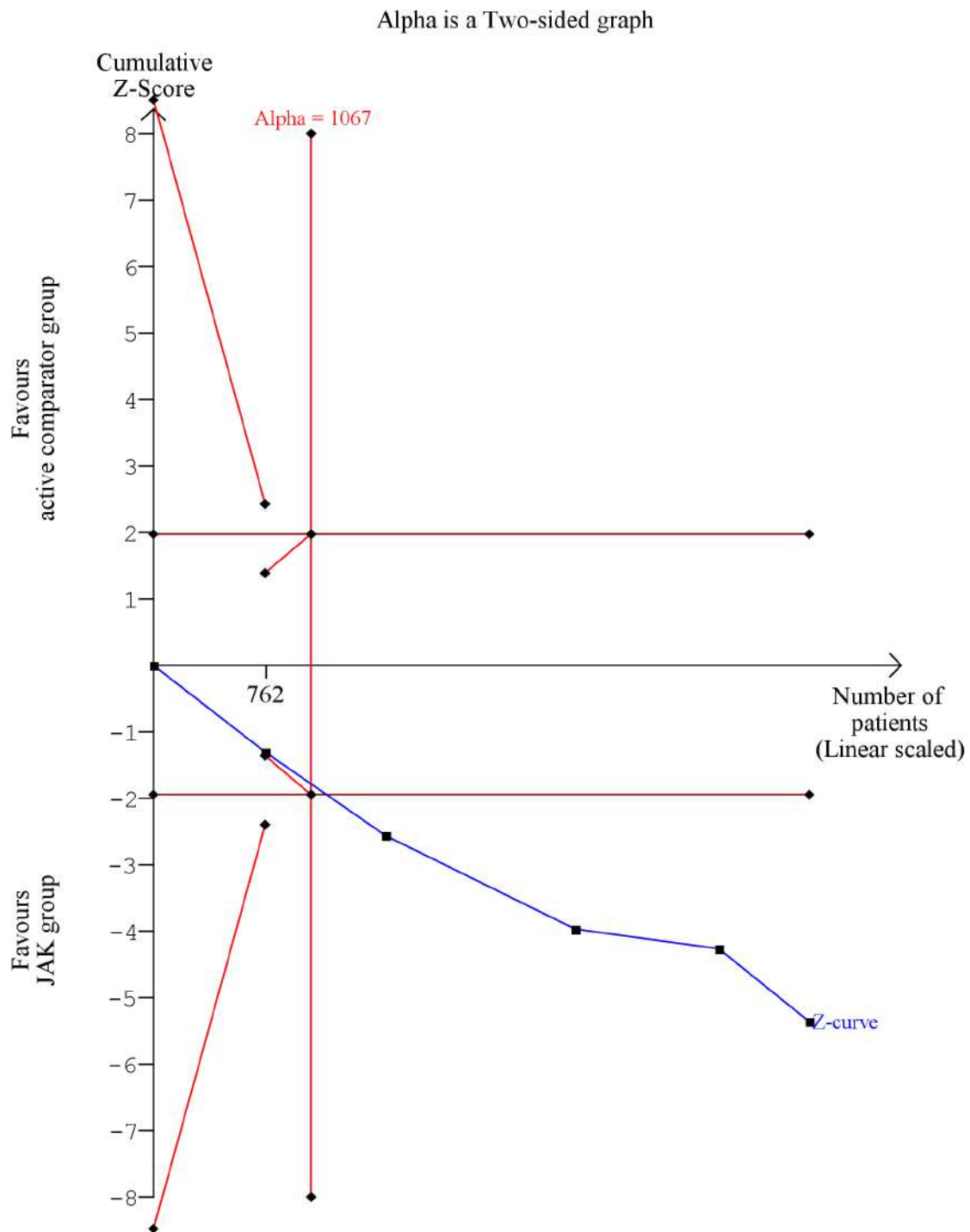
S169. Forest plot of studies comparing the LSM change from baseline in CRP value (mg/l) between patients treated with JAK inhibitors and bDMARDs within 6 months (CRP)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
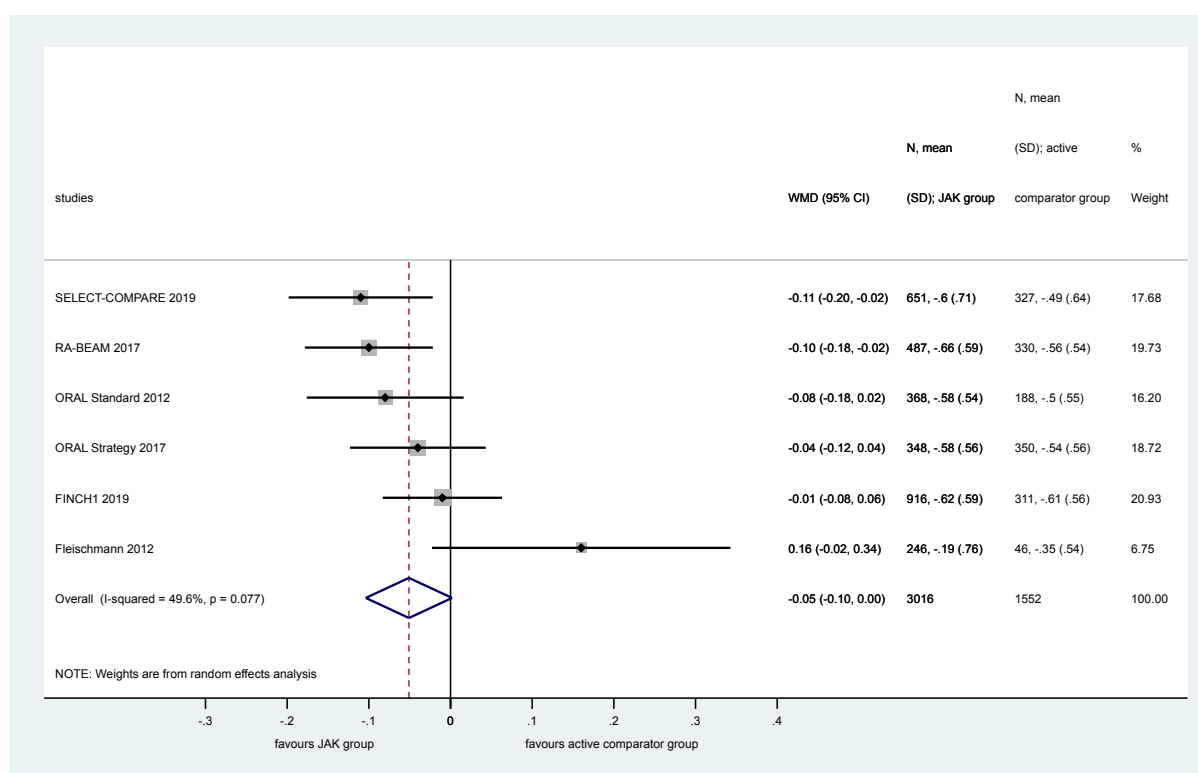
S170. Forest plot of studies comparing the LSM change from baseline in the score on HAQ-DI between patients treated with JAK inhibitors and bDMARDs within 6 months (HAQ-DI difference)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

S171. Forest plot of studies comparing the LSM change from baseline in the score on WPAI assessing overall work impairment (productivity loss) between patients treated with JAK inhibitors and bDMARDs within 6 months (WPAI OWI)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
SELECT-CHOICE 2020	-0.17 (-1.80, 1.46)	292, 5.63 (10.2)	294, 5.8 (9.98)	20.79
RA-BEAM 2017	0.12 (-1.25, 1.49)	487, 3.42 (10.1)	330, 3.3 (9.65)	29.25
ORAL Standard 2012	1.24 (-0.33, 2.81)	368, 4.62 (9.06)	188, 3.38 (8.91)	22.30
ORAL Strategy 2017	1.30 (-0.11, 2.71)	346, 5.7 (9.49)	349, 4.4 (9.53)	27.65
Overall (I-squared = 0.0%, p = 0.409)	0.64 (-0.11, 1.38)	1493	1161	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S172. Forest plot of studies comparing the LSM change from baseline in the score on SF-36 assessing the mental component score between patients treated with JAK inhibitors and bDMARDs within 6 months (SF-36 MCS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
ORAL Strategy 2017	0.10 (-1.09, 1.29)	346, 7.9 (8)	349, 7.8 (8.03)	16.41
FINCH1 2019	0.45 (-0.54, 1.44)	955, 8.85 (7.91)	325, 8.4 (7.89)	19.61
ORAL Standard 2012	1.13 (-0.12, 2.38)	368, 7.39 (7.13)	188, 6.26 (7.13)	15.53
RA-BEAM 2017	1.50 (0.44, 2.56)	487, 8.69 (7.69)	330, 7.19 (7.48)	18.52
SELECT-COMPARE 2019	1.60 (0.46, 2.74)	651, 7.9 (10.2)	327, 6.3 (7.63)	17.20
SELECT-CHOICE 2020	2.59 (1.11, 4.07)	292, 9.62 (9.29)	294, 7.03 (8.97)	12.73
Overall (I-squared = 47.6%, p = 0.090)	1.16 (0.50, 1.82)	3099	1813	100.00

NOTE: Weights are from random effects analysis

favours active comparator group | favours JAK group

S173. Forest plot of studies comparing the LSM change from baseline in the score on SF-36 assessing the physical component score between patients treated with JAK inhibitors and bDMARDs within 6 months (SF-36 PCS)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.

studies	WMD (95% CI)	N, mean (SD); JAK group	N, mean (SD); active comparator group	% Weight
FINCH1 2019	0.35 (-0.87, 1.57)	907, 9.15 (9.98)	304, 8.8 (9.19)	21.61
RA-BEAM 2017	0.39 (-0.86, 1.64)	487, 9.09 (9.04)	330, 8.7 (8.82)	20.76
SELECT-CHOICE 2020	1.26 (-0.49, 3.01)	290, 9.61 (10.9)	293, 8.35 (10.6)	10.54
ORAL Standard 2012	1.32 (-0.08, 2.72)	368, 6.36 (8.01)	188, 5.04 (7.95)	16.41
ORAL Strategy 2017	1.52 (0.13, 2.91)	348, 7.59 (9.33)	352, 6.07 (9.38)	16.75
SELECT-COMPARE 2019	1.60 (0.08, 3.12)	651, 9 (12.6)	327, 7.4 (10.8)	13.95
Overall (I-squared = 0.0%, p = 0.633)	0.98 (0.42, 1.55)	3051	1794	100.00

NOTE: Weights are from random effects analysis

favours active comparator group   favours JAK group

S174. Forest plot of studies comparing the LSM change from baseline in the score on FACIT-F between patients treated with JAK inhibitors and bDMARDs within 6 months (FACIT-F)

Black diamonds exhibit the individual effect of each study, whereas outer lines show the confidence intervals (CI), and the size of grey squares are proportional to the weight of each study. The blue diamond demonstrates the overall effect, the edges of which represent the CIs.
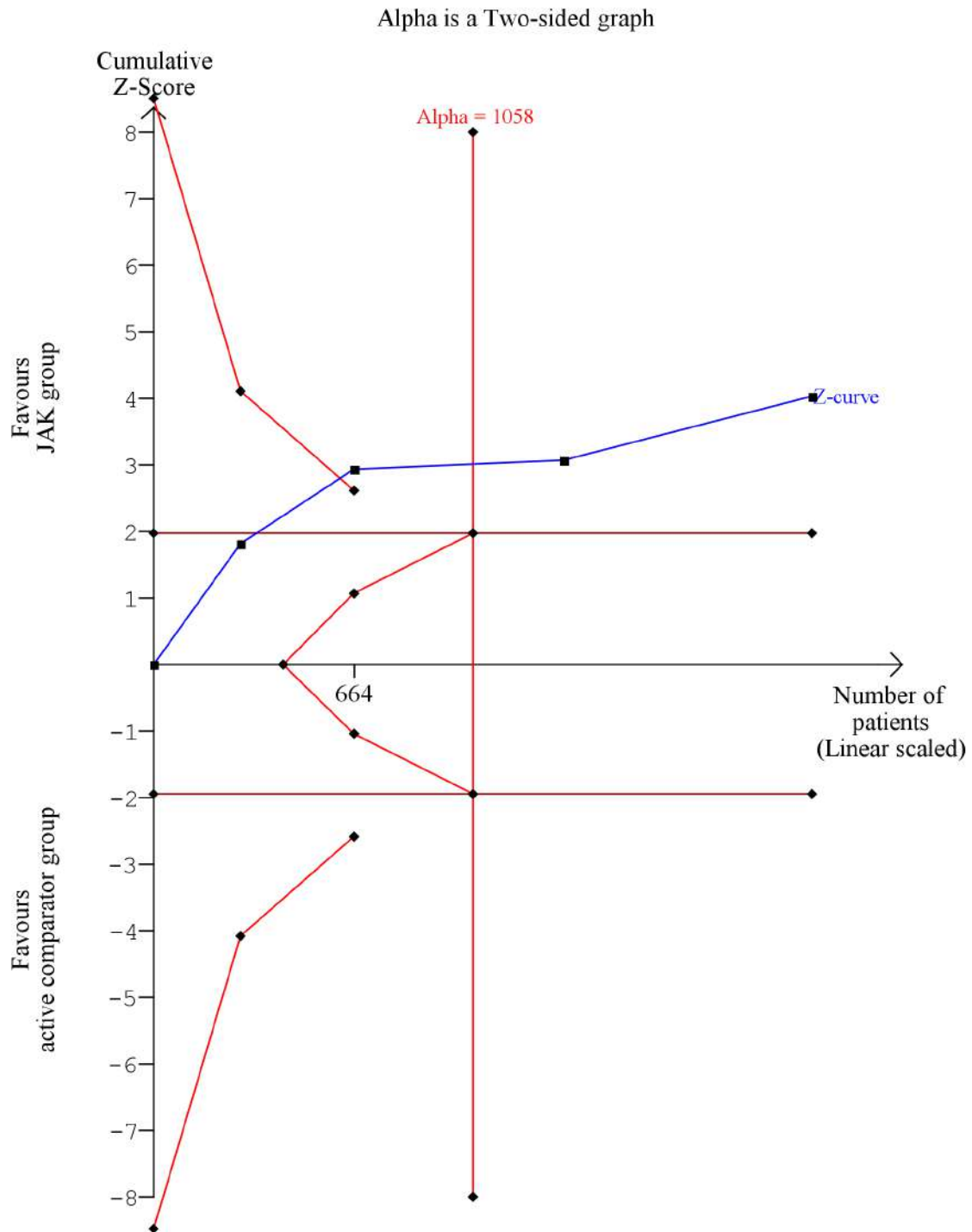
Alpha is a Two-sided graph

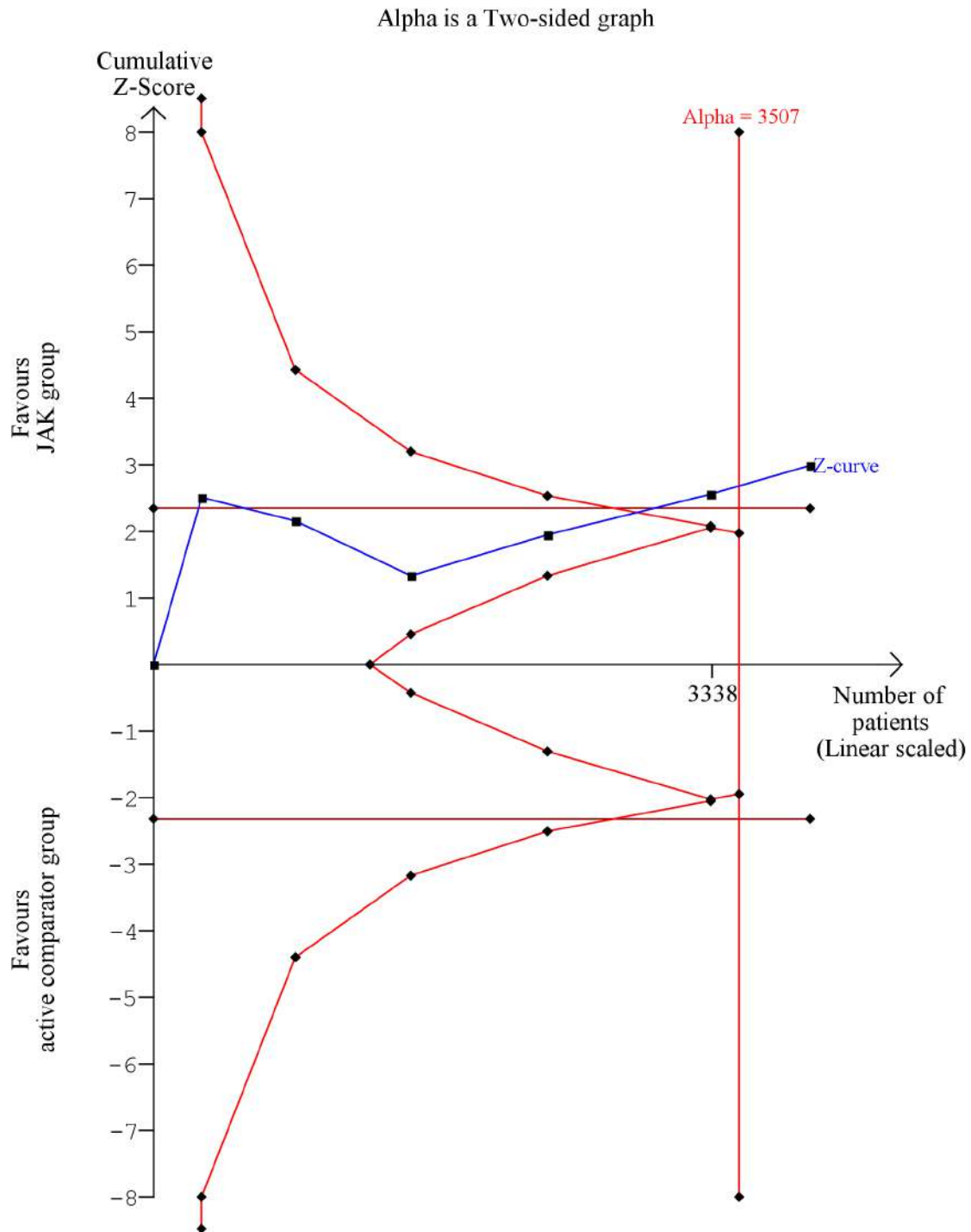**S175. TSA of ACR20 response (JAK-inhibitors vs bDMARDs)**

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.

Alpha is a Two-sided graph

S176. TSA of ACR50 response (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this results.

Alpha is a Two-sided graph

S177. TSA of ACR70 response (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this result
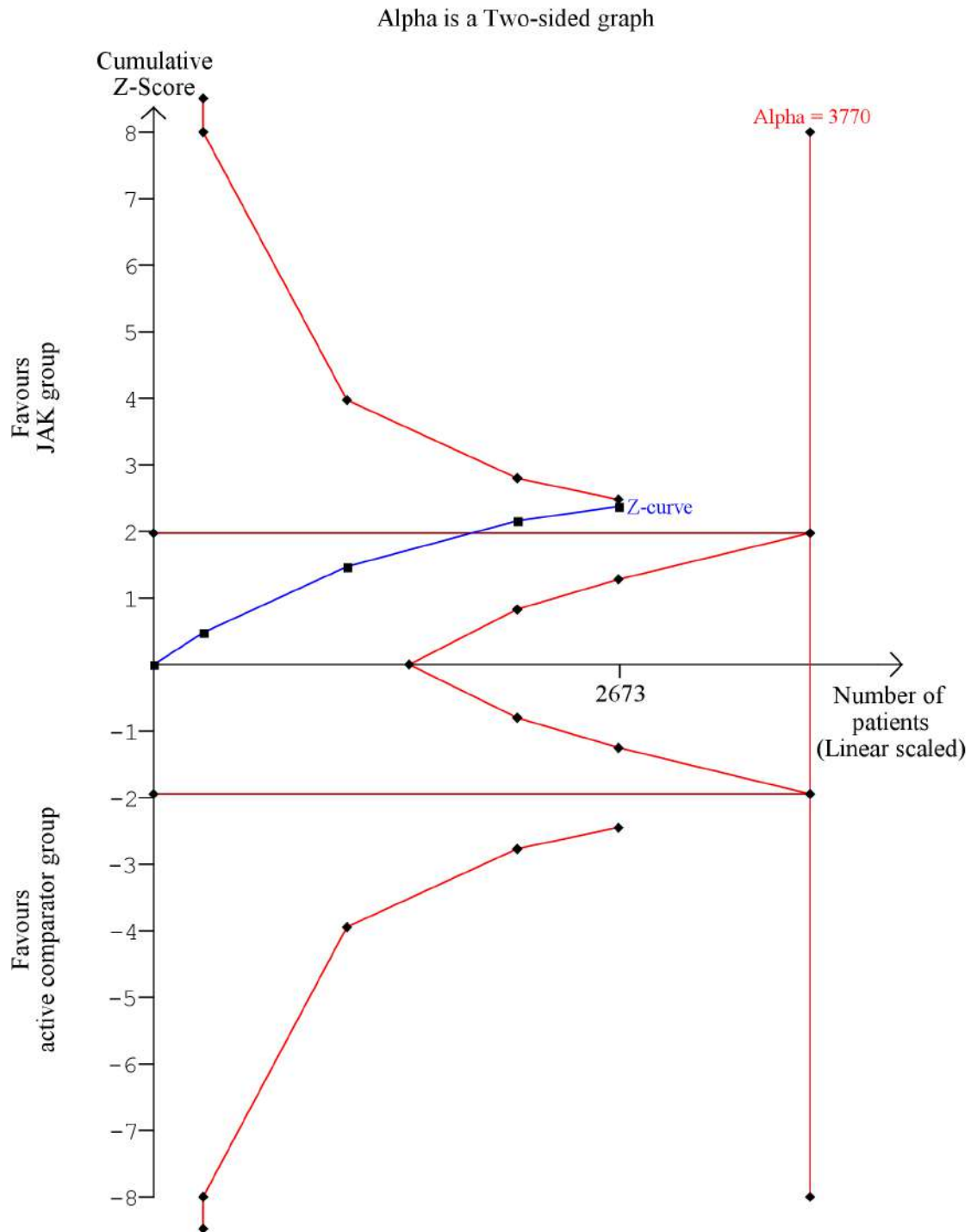
S178. TSA of CDAI remission (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.
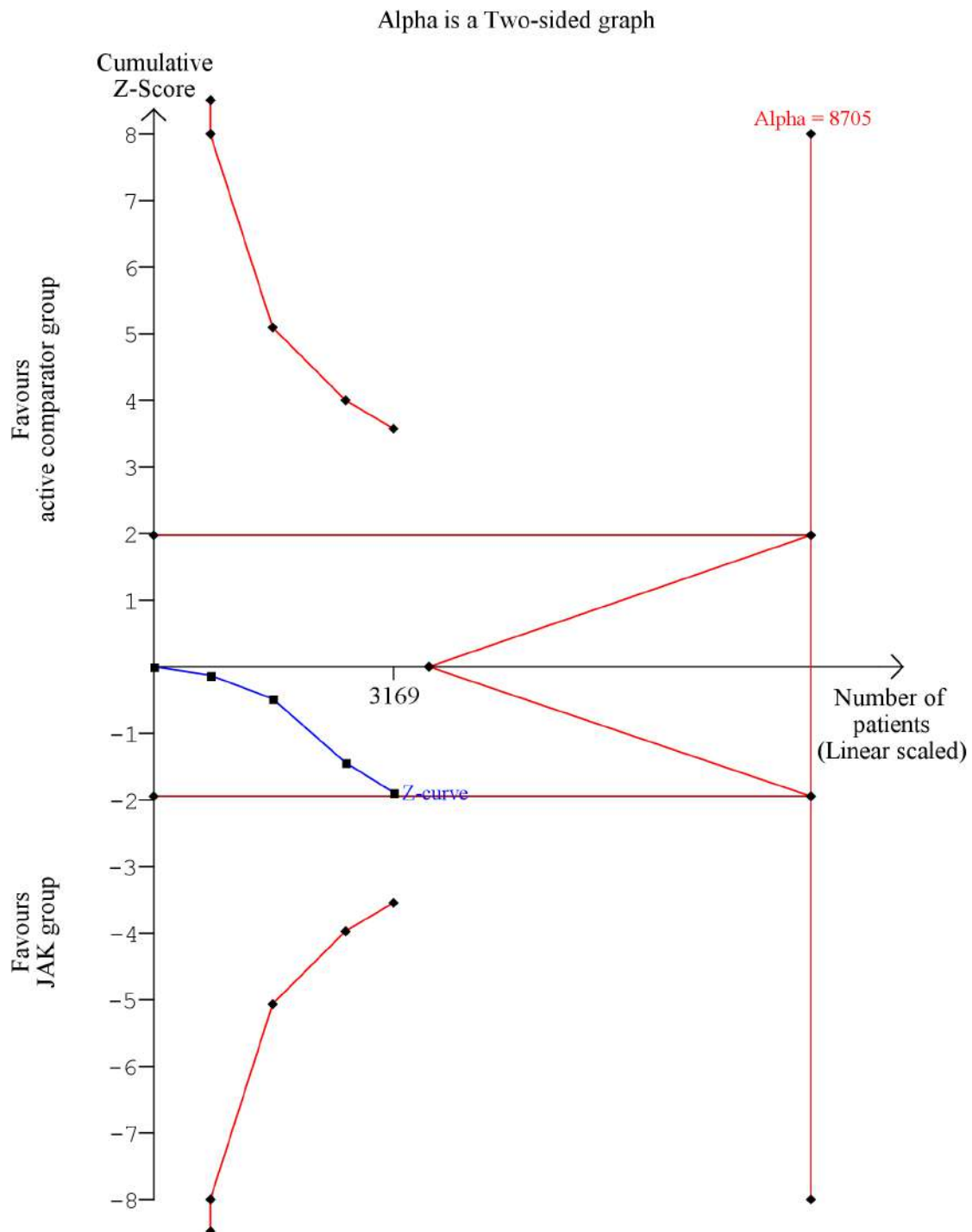
S179. TSA of SDAI remission (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.
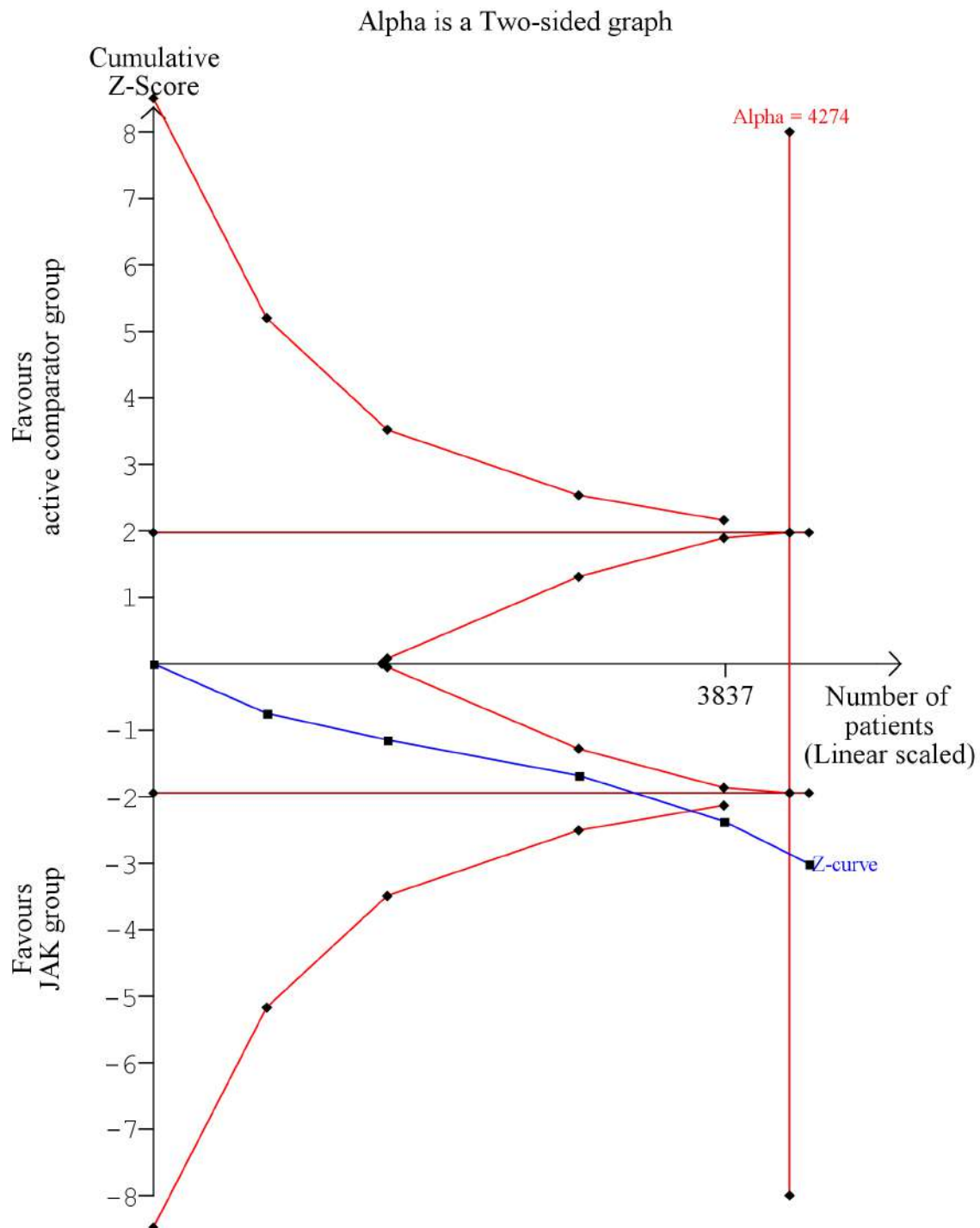
S180. TSA of DAS28-CRP remission within 6 months (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
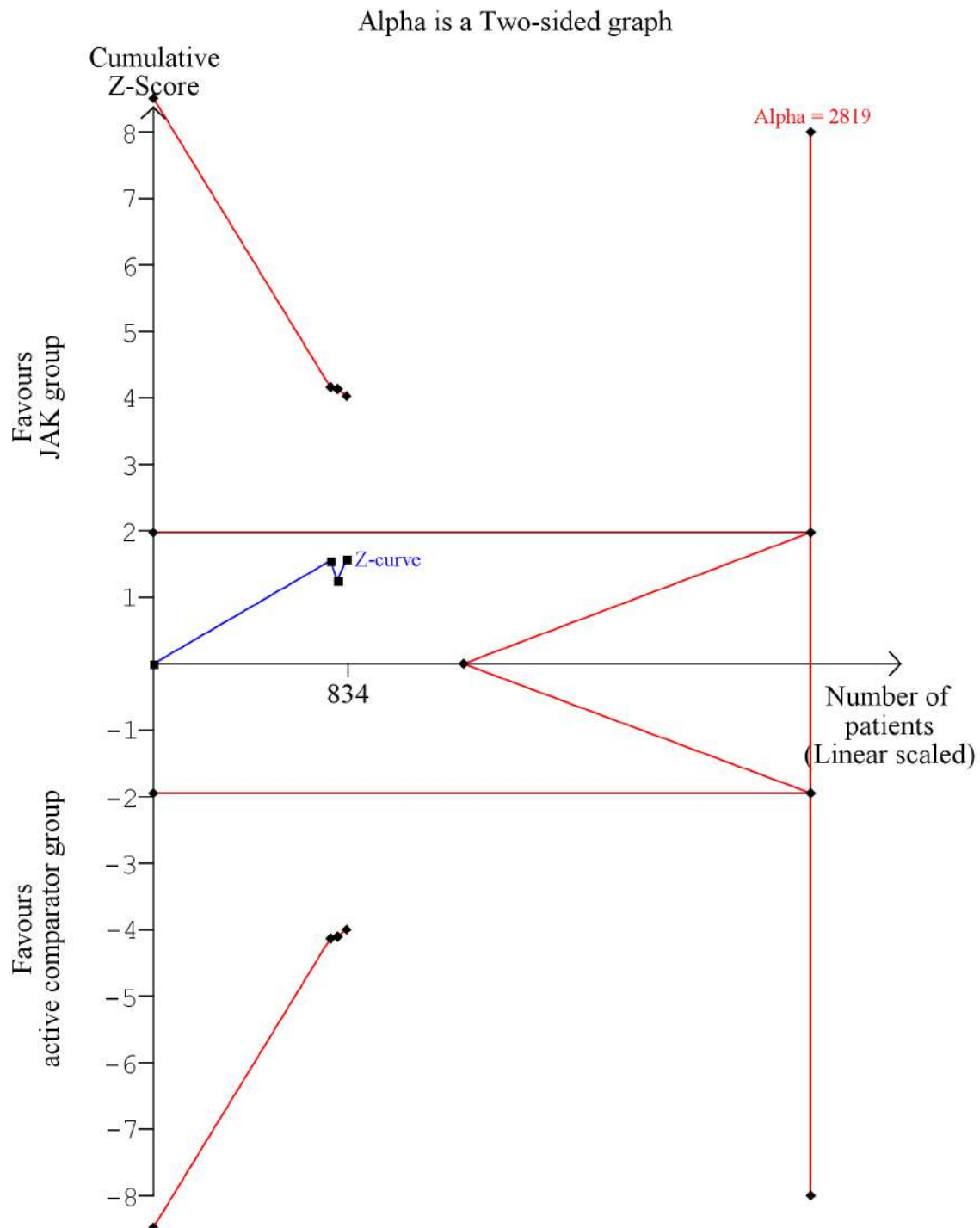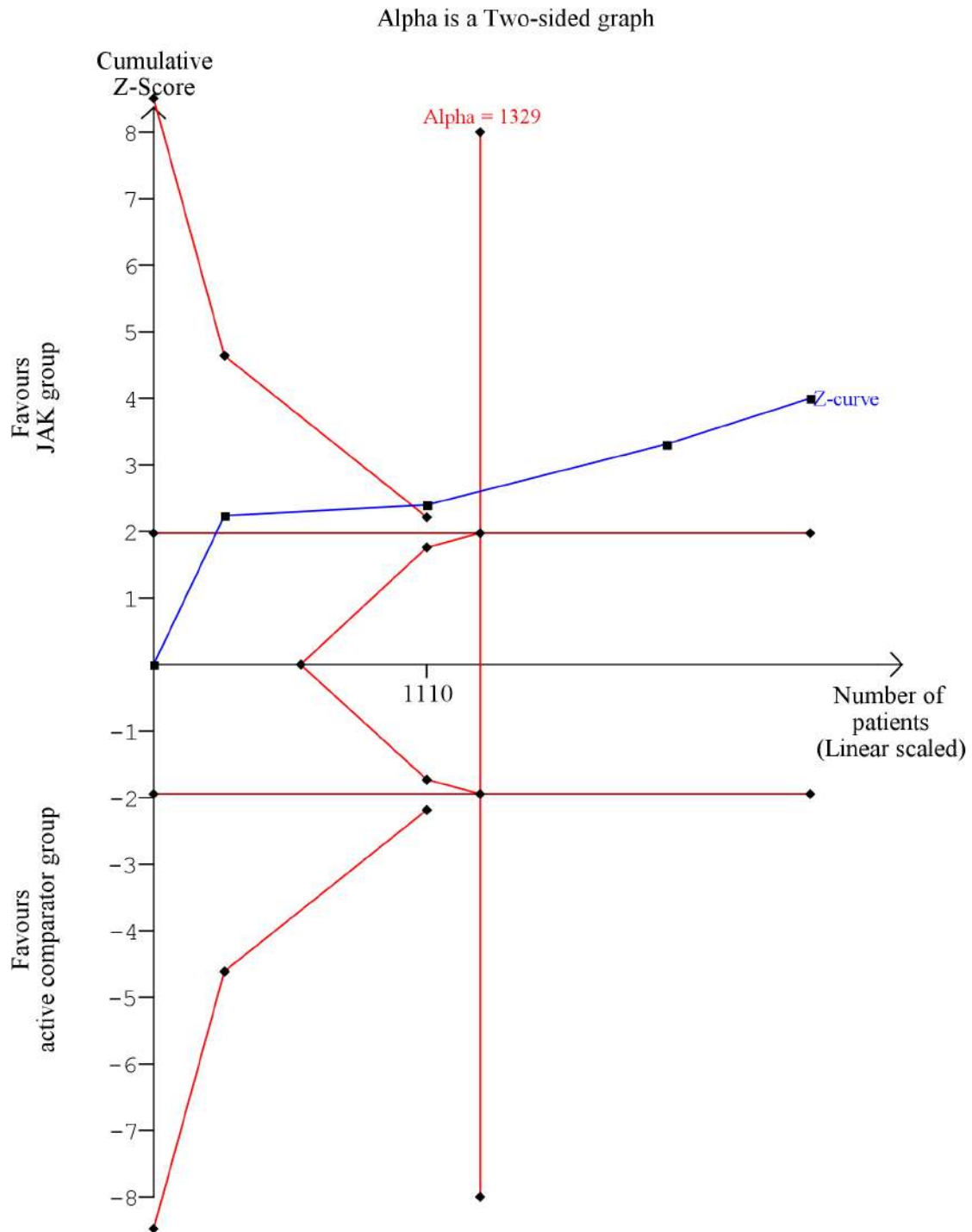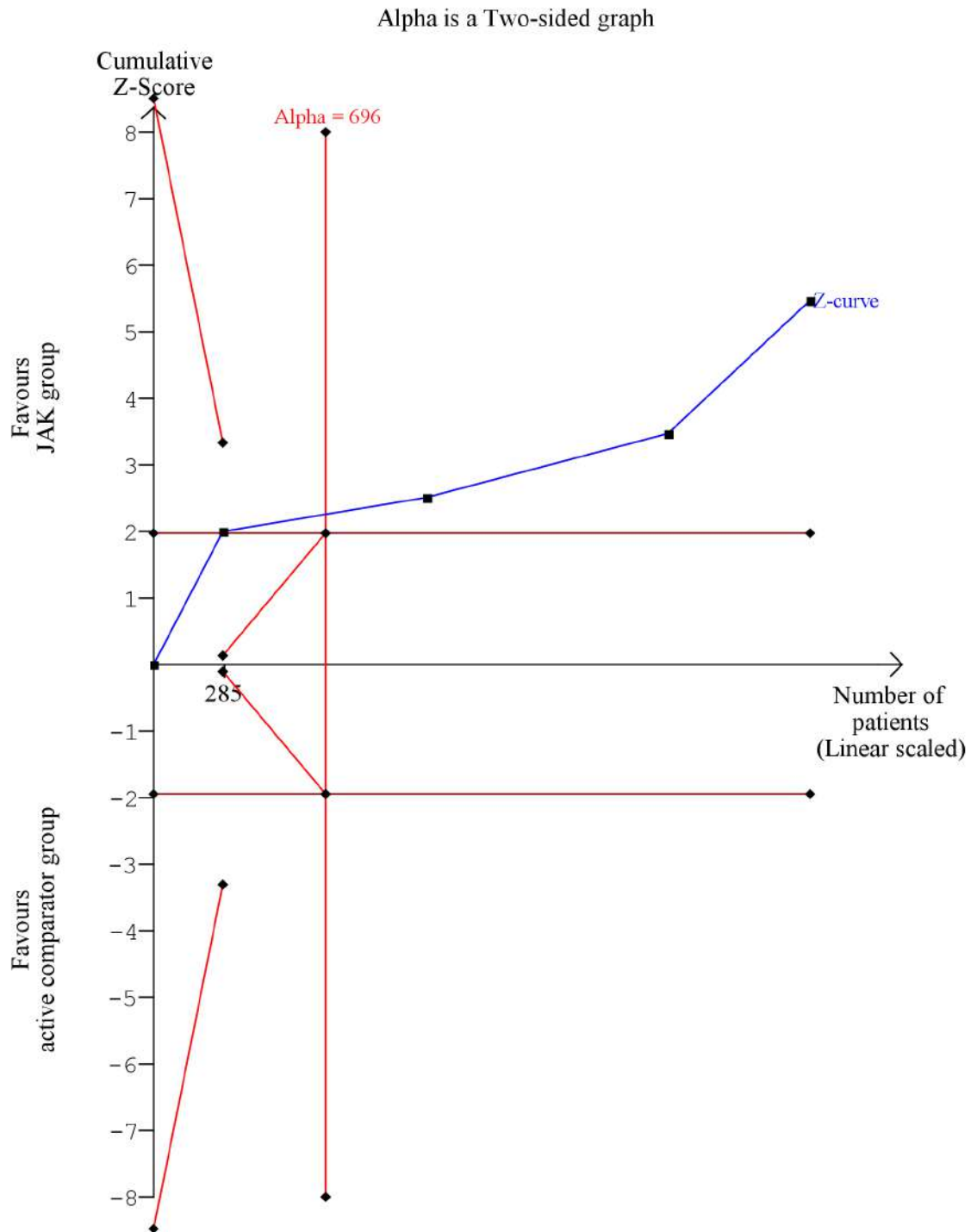
S181. TSA of DAS28-CRP remission over 6 months (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.

Alpha is a Two-sided graph

S182. TSA of DAS28-ESR remission (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.

Alpha is a Two-sided graph

S183. TSA of serious side effects (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.

Alpha is a Two-sided graph

S184. TSA of Pain VAS (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
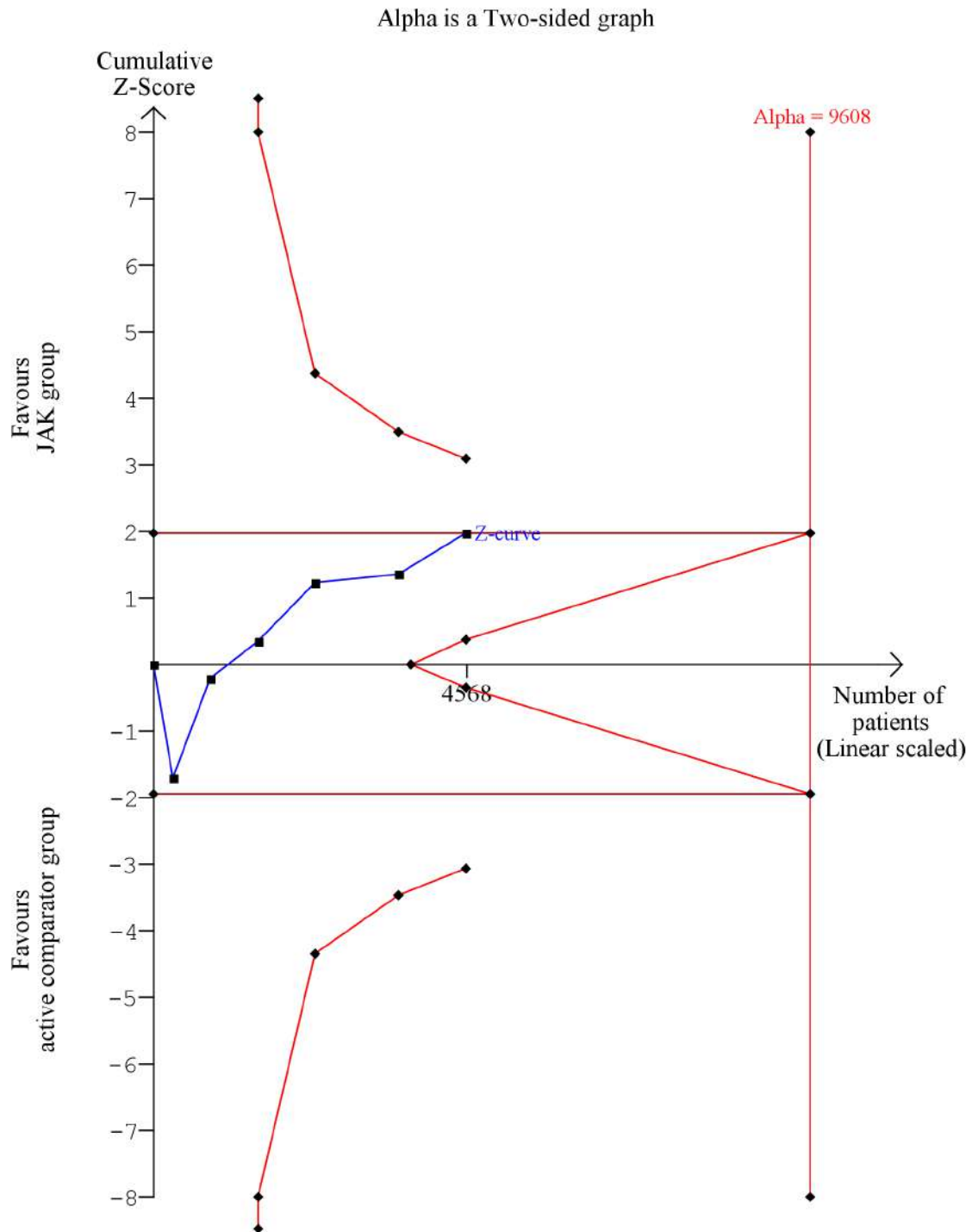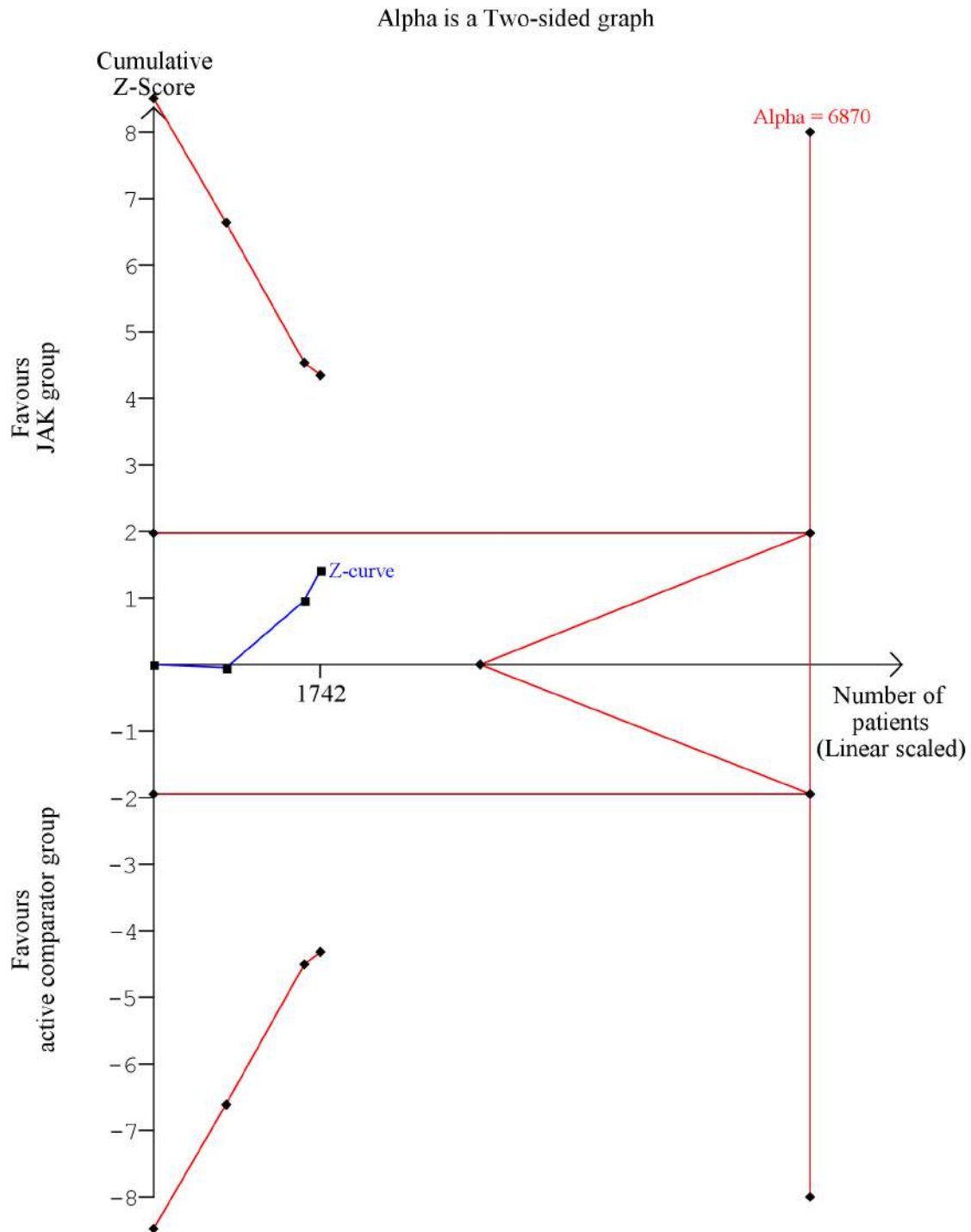
S185. TSA of PGA VAS (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

S186. TSA of PtGA VAS (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.
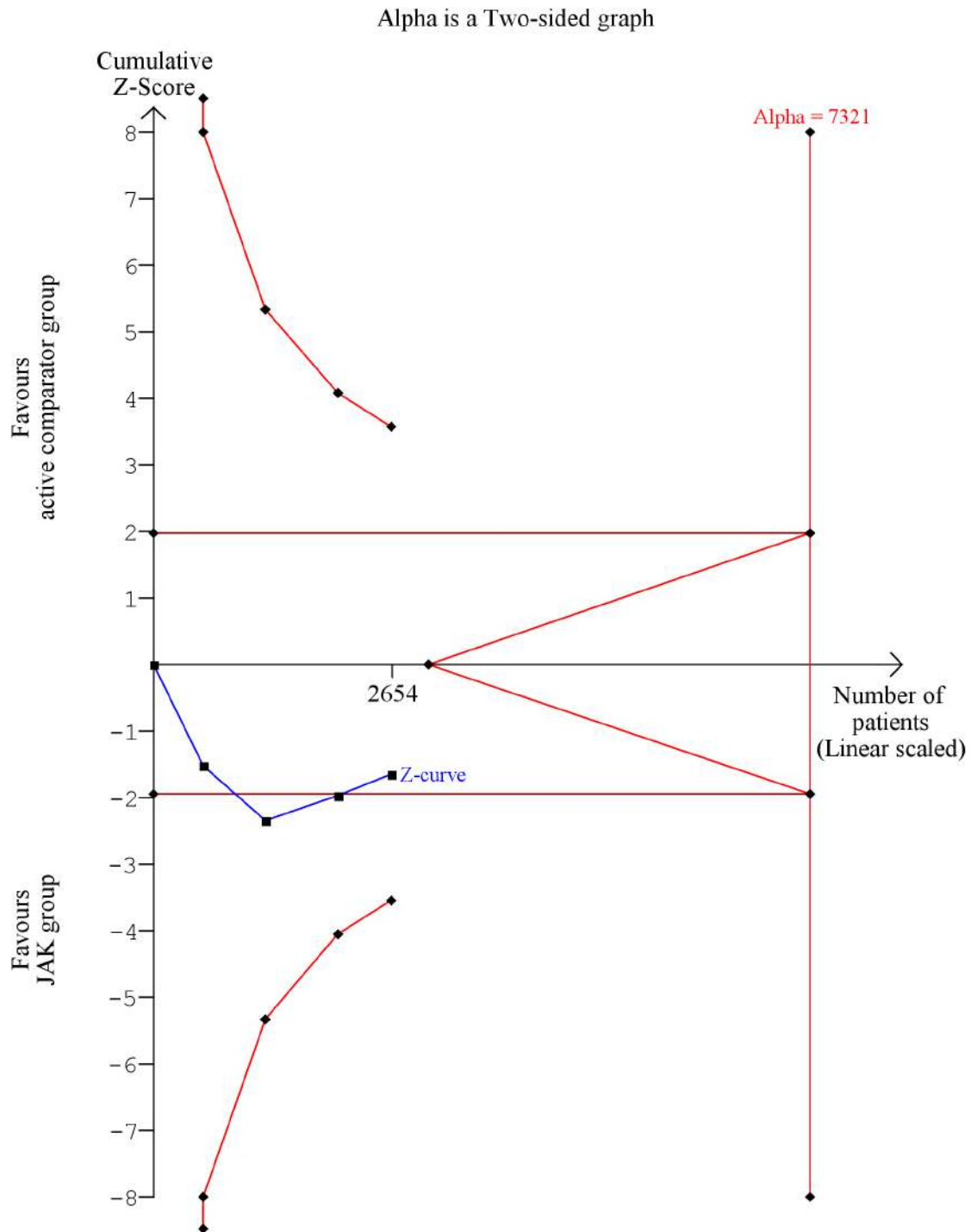
S187. TSA of Swollen Joint Count (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) but does not reach the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome but the addition of further studies might influence these results.
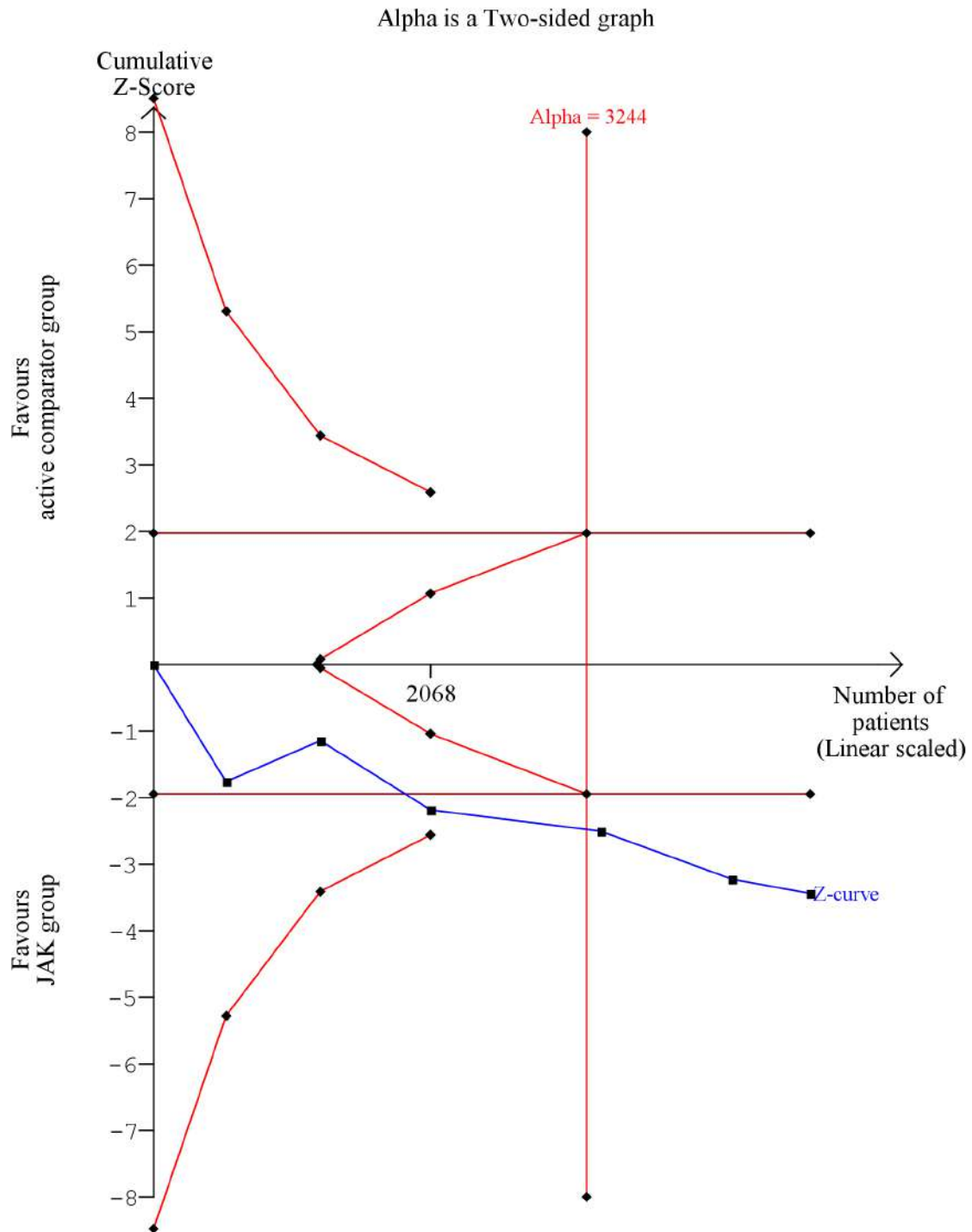
Alpha is a Two-sided graph

S188. TSA of Tender Joint Count (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph

S189. TSA of CRP (mg/l) (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence this result.

S190. TSA of HAQ-DI difference (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.
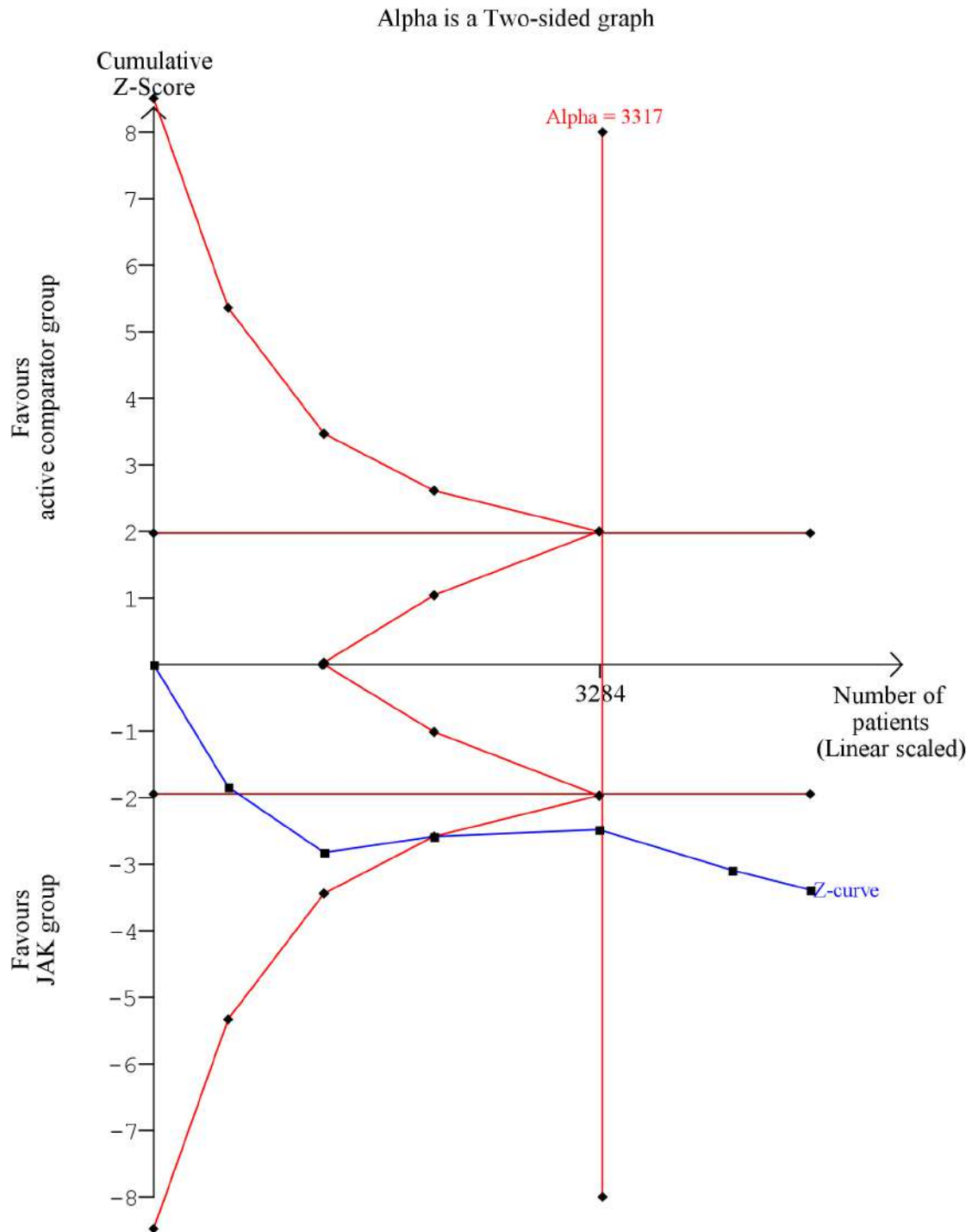
S191. TSA of WPAI OWI (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.
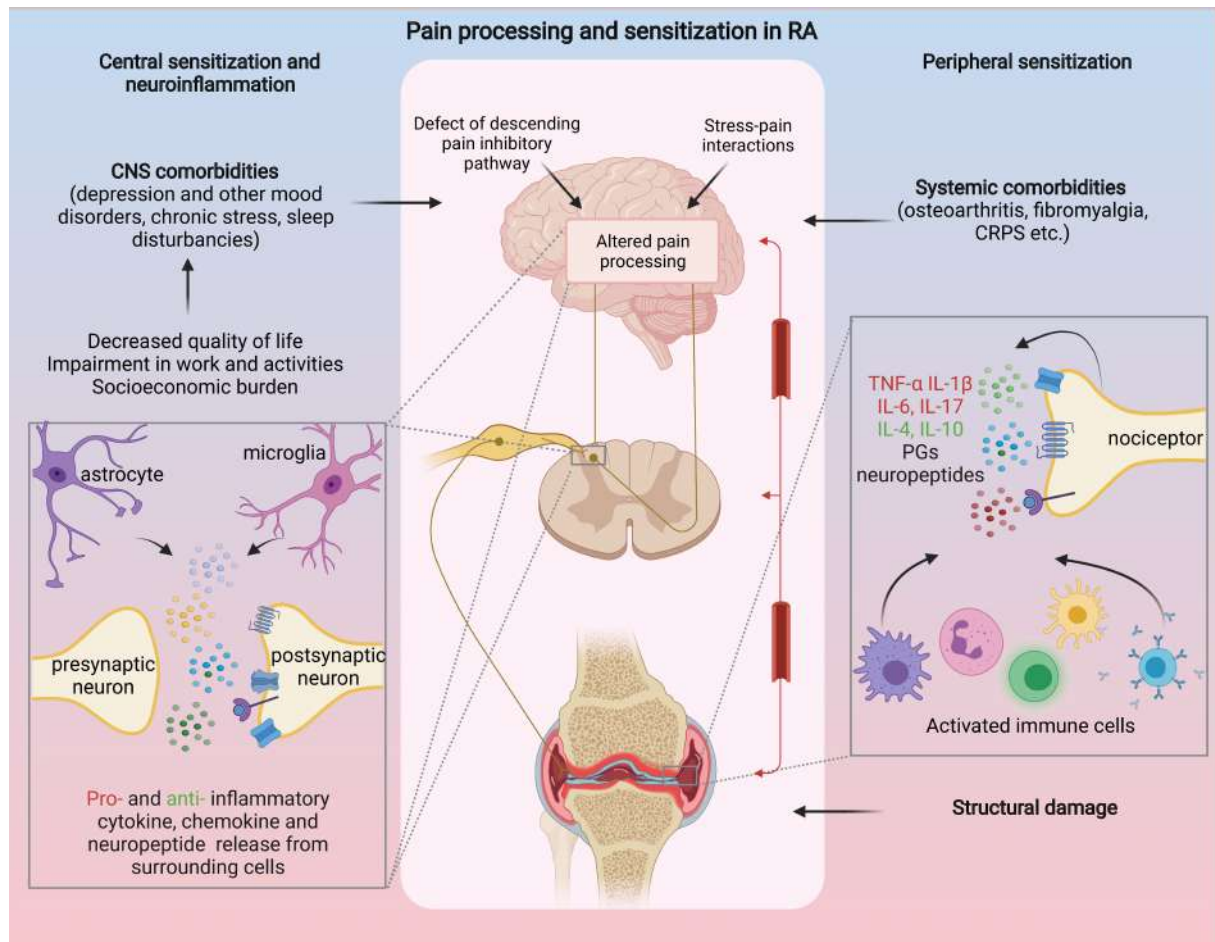
S192. TSA of SF-36 MCS (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) does not cross the conventional boundary (brown straight line) and does not reach the line of required sample size, therefore showing no statistically significant difference between groups in this outcome and the addition of further studies might influence these results.

Alpha is a Two-sided graph

S193. TSA of SF-36 PCS (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

Alpha is a Two-sided graph

S194. TSA of FACIT-F (JAK-inhibitors vs bDMARDs)

The blue line (Z-curve) crosses the conventional boundary (brown straight line) and trial sequential boundary (red lines sloping inwards), and reaches the line of required sample size, therefore showing the statistically significant superiority of JAK inhibitors in this outcome and the addition of further studies would not influence these results.

S195. Hypothesized inflammatory and non-inflammatory factors contributing to induce and maintain pain in RA.

Immune cells e.g. macrophages, neutrophil granulocytes, dendritic -, T- and B cells recruited in the joints produce several inflammatory cytokines (TNFa, IL-1b, IL-6, IL-17, IL-4, IL-10) that can activate their receptors on nociceptors leading to peripheral sensitization. Non-neuronal structures such as astrocytes and glial cells also participate in regulating inflammatory processes by producing cytokines, chemokines and neuropeptides involved in the pathogenesis of the disease. Neuroinflammatory processes also take place in the dorsal root ganglia and the central nervous system resulting in central sensitization. Locally activated immune cells and soluble molecules can also travel through the bloodstream contributing to the complex systemic effects. Together with the defect of the descending pain inhibitory routes, these changes lead to the alteration of pain processing. Chronic inflammation can result in irreversible structural damage and the presence of comorbidities can also play an important role contributing to pain. Mental disorders and the stress related to coping with the disease can affect pain modulation via the complex interactions between the regulation of stress and pain, whereas systemic diseases can also influence pain symptoms.

RA, rheumatoid arthritis; TNF-a, tumor necrosis factor-alpha; IL-1b, Interleukin-1 beta, IL-6, Interleukin-6; IL-17, Interleukin-17; IL-4, Interleukin-4, IL-10, Interleukin-10; PG, prostaglandin; CNS, central nervous system; CRPS, complex regional pain syndrome

Figure S195 was prepared with the help of BioRender (www.biorender.com)