SUPPLEMENTARY APPENDIX

Supplementary appendix 1: Selection criteria

Inclusion criteria included

- Active rheumatoid arthritis (RA) of ≥6 months' duration at baseline, Disease Activity Score using 28 joints and erythrocyte sedimentation rate (DAS28-ESR) of >3.2 and diagnosis according to the 1987 American College of Rheumatology (ACR) classification criteria
- Able and willing to give written informed consent to participate and comply with the requirements of the study and to give written consent for data protection (legal requirement in Germany: Datenschutzrechtliche Einwilligung)
- Able to receive treatment on an outpatient basis
- Age ≥18 years
- ESR ≥28 mm/h or C-reactive protein ≥0.7 mg/dl (SI, ≥7 mg/L) at screening and/or baseline
- Receiving one or more permitted disease-modifying anti-rheumatic drugs (DMARDs); current DMARD therapy must have been at a stable dose for ≥4 weeks before baseline
- Oral corticosteroids (≤10 mg/day prednisone or equivalent) permitted if dose was stable for ≥4 weeks before baseline
- Women of childbearing potential allowed to participate in this trial only if using reliable, highly effective contraceptives (allowed methods of birth control, *i.e.* with a failure rate of <1% per year, are implants, injectables, combined oral contraceptives, intrauterine systems [only hormonal contraceptive coil], sexual abstinence or partner who has undergone vasectomy)
- Women of childbearing potential or <1 year after menopause (unless surgically sterile) must have a negative pregnancy test result (urine β-human chorionic gonadotropin) at screening. Pregnancy test must be repeated at baseline in case of absence of menstruation or irregular menstrual cycle

Exclusion criteria included

Related to general health

- Major surgery (including joint surgery) ≤8 weeks before baseline or planned major surgery within the study duration (≤32 weeks after baseline)
- Functional class IV as identified by the ACR classification of Functional Status in Rheumatoid Arthritis
- Rheumatic autoimmune disease other than RA, including systemic lupus erythematosus, mixed connective tissue disease, scleroderma, polymyositis or significant systemic involvement secondary to RA (*e.g.* vasculitis, pulmonary fibrosis or Felty's syndrome). Sjögren's syndrome with RA was allowed
- History of or current inflammatory joint disease other than RA (*e.g.* gout, reactive arthritis, psoriatic arthritis, seronegative spondyloarthropathy, active Lyme disease) (active Lyme disease was defined by typical clinical signs and symptoms and presence of anti-Borrelia-specific immunoglobulin M identified in two independent test systems, such as enzyme-linked immunosorbent assay and immunoblot analysis; lymphocyte transformation test was not approved as a test system)

Related to medications

- Current participation in another clinical trial or past participation in another clinical trial within 30 days before screening (or five half-lives of the Investigational Medicinal Product, whichever is longer) or patients who previously participated in this trial.
- Previous treatment with any cell-depleting therapies, including investigational agents (*e.g.* CAMPATH, anti-CD4, anti-CD5, anti-CD3, anti-CD19 and anti-CD20)
- Treatment with intravenous gamma globulin plasmapheresis or immunosorbent column within 6 months before baseline
- Intra-articular or parenteral corticosteroids within 4 weeks before baseline; injection of intra-articular steroids while on study medication discouraged but may be used in a limited fashion
- Immunisation with a live/attenuated vaccine within 4 weeks before baseline
- Previous treatment with tocilizumab, rituximab or other biologic DMARDs such as tumour necrosis factor blockers
- Combination therapy with methotrexate and leflunomide within 4 weeks before baseline
- Any previous treatment with alkylating agents, such as cyclophosphamide or chlorambucil, or with total lymphoid irradiation

Related to general safety

- Severe heart failure (New York Heart Association class IV) or severe, uncontrolled cardiac disease
- History of severe allergic or anaphylactic reaction to human, humanised or murine monoclonal antibodies
- Evidence of serious uncontrolled concomitant cardiovascular, nervous system, pulmonary (including obstructive pulmonary disease), renal, hepatic, endocrine (including uncontrolled diabetes mellitus) or gastrointestinal disease
- Uncontrolled disease states, such as asthma, psoriasis or inflammatory bowel disease in which flares are commonly treated with oral or parenteral corticosteroids
- History of diverticulitis, diverticulosis requiring antibiotic treatment or chronic ulcerative lower gastrointestinal (GI) disease such as Crohn's disease, ulcerative colitis or other symptomatic lower GI conditions that might predispose to perforations OR evidence of serious uncontrolled concomitant gastrointestinal disease
- Current liver disease as determined by principal investigator (patients with history of alanine aminotransferase [ALT] elevation will not be excluded)
- Known history of or active recurrent bacterial, viral, fungal, mycobacterial or other infection (including but not limited to tuberculosis and atypical mycobacterial disease, granulomatous disease on chest x-ray [x-ray should not be older than 90 days related to treatment start], hepatitis B and C and herpes zoster but excluding fungal infections of nail beds) or any major episode of infection requiring hospitalisation or treatment with intravenous antibiotics ≤4 weeks before screening or oral antibiotics ≤2 weeks before screening
- Active tuberculosis requiring treatment within the previous 3 years. QuantiFERON (Qiagen, Hilden, Germany) testing (interferon gamma release assay [IGRA]) should be used rather than protein derivative tuberculin skin (PPD) or ELISPOT testing, and results must be negative. Results on chest x-ray also must be negative
  - Exception: If a PPD test result was positive, treatment for latent tuberculosis must be completed and subsequent result on chest x-ray must be negative. In addition, retesting with IGRA or ELISPOT must be performed, and the result must be negative
  - Primary or secondary immunodeficiency (history of or active)
- History of malignancy, including solid tumours and haemato-

logical malignancies (except basal cell carcinoma of the skin that has been excised and cured)
- Pregnant or nursing (breastfeeding) women
- History of alcohol, drug or chemical abuse ≤6 months before screening
- Neuropathies or other painful conditions (only in reference to joint status) that might interfere with pain evaluation
- Any severe neurological diseases and any history of or current demyelinating diseases
- Patients with lack of peripheral venous access
- Body weight >130 kg

Related to laboratory findings
- Serum creatinine >1.4 mg/dl (SI, >124 μmol/L) in female patients and >1.6 mg/dl (SI, >141μmol/L) in male patients
- ALT or aspartate aminotransferase (AST) >1.5× upper limit of normal (ULN) (if initial sample yields ALT or AST >1.5× ULN, a second sample may be taken and tested during the screening period)
- Platelet count <100,000/mm^3 (SI, <100 × 10^9/L)
- Haemoglobin <8.5 g/dl (SI, <5.3 mmol/L)
- White blood cell count <3000/mm^3 (SI, <3.0 × 10^9/L)
- Absolute neutrophil count <2000/mm^3 (SI, <2.0 × 10^9/L)
- Absolute lymphocyte count <500/mm^3 (SI, <0.5 ×10^9/L)
- Positive hepatitis B surface antigen or hepatitis C antibody
- Total bilirubin >ULN (if initial sample yields bilirubin >ULN, a second sample may be taken and tested during the screening period)
- Triglycerides >900 mg/dl (SI, >10 mmol/L) at screening (non-fasted) or low-density lipoprotein >160 mg/dl (SI, >4.1 mmol/L)

Related to formal aspects
- Age <18 years, inability to understand the aim, importance and consequences of the study and inability to give legal informed consent
- History of a psychological illness or a condition that could interfere with the ability to understand the requirements of the study
- Institutionalisation because of regulatory or judicial order
- Possible dependence on the sponsor or the investigator

Supplementary appendix 2: Additional statistical methods

Based on the assumptions for the primary efficacy endpoint, a sample size of 428 patients was estimated for the first treatment period to provide 80% power to detect DAS28 remission at a rate of >45%. Patients who initiated escape therapy or withdrew from the study or for whom DAS28 could not be determined were considered non-responders (applicable for all binary response parameters). With expected drop-out rates between 10 to 15% a sample size of n=500 was chosen.

The ITT1 subpopulation included week 16 early tocilizumab responders (patients who received ≥1 dose of tocilizumab in the first treatment period and completed the study, achieving DAS28-ESR remission, at week 16). The ITT2 subpopulation was defined to include week 16 partial tocilizumab responders receiving continued treatment with tocilizumab (patients who received ≥1 dose of tocilizumab in the first treatment period with ≥1 efficacy measurement under TCZ, receiving TCZ in the second treatment period). The ITT3 subpopulation included week 16 tocilizumab non-responders (patients who received ≥1 dose of tocilizumab in the first period, ≥1 dose of rituximab in the second period, and had ≥1 efficacy measurement during rituximab treatment).

To examine the patient profile of early *versus* late tocilizumab responders in more depth and to identify potential predictive factors for tocilizumab response, post hoc definitions were assigned for early responders (ITT1), late responders (ITT2 patients who achieved DAS28 <2.6 at week 32 and did not discontinue before week 32), EULAR good responders (ITT1 and ITT2 patients with good EULAR response at week 16 and/or 32) and EULAR non-responders (EULAR non-responders in ITT3 at week 16 and in ITT2 at week 32). Logistic regression analyses were performed using these post hoc definitions to determine the effect of various baseline parameters (age, sex, body mass index [BMI], presence of any comorbid disease, disease duration, investigator's and patient's assessments of disease activity [VAS], patient's assessment of pain, TJC28, SJC28, ESR, rheumatoid factor [RF], anti-cyclic citrullinated peptide [CCP] antibodies, CRP, white blood cell [WBC] count, lymphocyte count and Health Assessment Questionnaire-Disability Index [HAQ-DI]) on the likelihood of early or late response. EULAR response was also evaluated to further define responders and non-responders. In this post hoc analysis, patients who achieved good EULAR response (week 16 for early responders [ITT1] and week 32 for late responders [ITT2]) were compared with EULAR non-responders (week 32 for late responders [ITT2] and week 16 for patients who switched to rituximab [ITT3]). A two-step approach was used: the first model included all variables and the second model included variables that were significant (p<0.1) in the first model.

Supplementary appendix 3: Concomitant csDMARD and corticosteroid treatment

At baseline, the median daily corticosteroid dose was 5 mg and the median weekly methotrexate dose was 15 mg in the ITT1 and ITT2 subpopulations. No change in corticosteroid dose was reported for 89% of ITT1 and 85.1% of ITT2 patients; dose reductions were documented in 9.0% of ITT1 and 12.8% of ITT2 patients, and dose increases were documented in 1.9% and 2.0%, respectively. No change in methotrexate dose was reported for 94.8% of ITT1 patients and 88.3% of ITT2 patients; dose reductions were reported for 5.2% of ITT1 and 10.8% of ITT2 patients, and dose increase was reported in only one ITT2 patient. In the ITT3 subpopulation, the median daily corticosteroid dose was 5 mg and the median weekly methotrexate dose was 20 mg. Corticosteroid dose was unchanged through week 32 in 88.9% of ITT3 patients; 11.1% had their dose decreased, and no patients had their dose increased. Methotrexate dose was unchanged in 76.2% of ITT3 patients; 14.3% had their dose decreased, and 9.5% had their dose increased through week 32.

Supplementary appendix 4: Additional safety results

Most AEs were mild or moderate in severity, and the types of SAEs reported were diverse. There did not appear to be any pattern in the type of SAEs reported in patients who received subsequent rituximab; individual SAEs were reported in eight patients. Aside from infections, other SAEs reported during the study in tocilizumab-treated patients included serious anaphylaxis/hypersensitivity reactions (four patients), malignancies (one patient with malignant melanoma and one with laryngeal squamous cell carcinoma), myocardial infarction (two patients; one subsequently received rituximab), serious bleeding (two patients; neither event was considered an AE of special interest), serious diverticulitis (two patients) and gastrointestinal perforation (one patient). No stroke, tuberculosis or demyelinating disorders were reported. To account for the differences in person-time at risk (236.8 patient-years [PY] in the 490 patients in the tocilizumab safety population and 35.5 PY in the 27 patients in the rituximab safety population), SAE rates were calculated per 100 PY. Rates of serious cardiac disorders, musculoskeletal and connective tissue disorders and general disorders and administration site conditions were higher in patients who switched from tocilizumab to rituximab, whereas rates of serious infections and infestations and injury, poisoning and procedural complications were higher in patients who received only tocilizumab.

Of the total 519 patients, 333 and 78 were still receiving concomitant methotrexate and leflunomide at their last study visit. Early discontinuation from the study was higher in patients receiving concomitant leflunomide than in those with concomitant methotrexate (23.1% *vs*. 10.2%), as was the number of previous csDMARDs (≥2 csDMARDs: 80.8% *vs*. 50.8%), whereas the proportion in DAS28 remission at week 16 was lower (30.8% *vs*. 46.8%). Incidences of adverse events (65.4% *vs*. 60.7%) and serious infections (5.1% *vs*. 1.8%) were also higher in patients with concomitant leflunomide, and incidences of hepatic events were similar.

Supplementary Table SI. Baseline demographics and disease characteristics in the subpopulations.

	ITT1 n=217	ITT2 n=213	ITT3 n=27
Female, n (%)	141 (65.0)	151 (70.9)	18 (66.7)
Age, years	54.0 (46.0; 62.0)	58.0 (49.0; 65.0)	56.0 (51.0; 62.0)
BMI, kg/m^2	26.7 (23.6; 30.1)	27.3 (23.8; 30.6)	27.2 (23.2; 34.3)
RA disease duration, years	4.8 (1.6; 10.6)	3.6 (1.5; 9.1)	6.8 (2.1; 8.0)
No. of previous csDMARDs*	2.0 (1.0; 2.0)	2.0 (1.0; 3.0)	2.0 (1.0; 3.0)
No. of concomitant csDMARDs, n (%)			
Any	210 (96.8)	210 (98.6)	25 (92.6)
1	191 (88.0)	199 (93.4)	23 (85.2)
2	17 (7.8)	9 (4.2)	1 (3.7)
3	2 (0.9)	2 (0.9)	1 (3.7)
Concomitant csDMARDs,† n (%)			
Methotrexate	160 (73.3)	146 (68.5)	22 (81.5)
Leflunomide	28 (12.9)	44 (20.7)	2 (7.4)
Sulphasalazine	16 (7.4)	18 (8.5)	2 (7.4)
(Hydroxy-)chloroquine	21 (9.7)	13 (6.1)	2 (7.4)
Receiving corticosteroids	164 (75.6)	174 (81.7)	21 (77.8)
Corticosteroid dose, mg/day			
Median (IQR)	5.0 (5.0; 7.5)	5.0 (5.0; 7.5)	5.0 (5.0; 7.5)
Mean (SD)	5.99 (2.48)	6.15 (2.62)	6.11 (2.21)
Methotrexate, mg/week			
Median (IQR)	15.0 (15.0; 20.0)	15.0 (15.0; 15.0)	20.0 (15.0; 20.0)
Mean (SD)	15.41 (4.33)	15.05 (4.04)	18.75 (4.55)
RF, U/mL	50.0 (20.0; 164.0)	49.3 (20.0; 201.3)	28.0 (10.8; 111.8)
Anti-CCP, U/mL	90.9 (7.0; 252.0)	47.0 (2.9; 200.0)	55.0 (2.1; 200.0)
RF positive, n/N (%)‡	137/198 (69.2)	144/198 (72.7)	17/26 (65.4)
Anti-CCP positive, n/N (%)‡	110/144 (76.4)	102/137 (74.5)	14/18 (77.8)
DAS28-ESR	5.3 (4.6; 6.1)	5.9 (5.4; 6.6)	5.7 (4.8; 6.4)
SJC28	6.0 (4.0; 9.0)	8.0 (6.0; 12.0)	7.0 (4.0; 16.0)
TJC28	7.0 (5.0; 12.0)	11.0 (7.0; 15.0)	11.0 (8.0; 17.0)
ESR, mm/h	29.0 (20.0; 44.0)	34.0 (25.0; 42.0)	22.0 (14.0; 40.0)
CRP, mg/L	11.0 (6.0; 20.1)	9.5 (5.0; 18.5)	9.8 (5.3; 15.8)
Patient DA VAS, mm	60.0 (47.0; 77.0)	71.0 (55.0; 83.0)	66.0 (56.0; 74.0)
SJC66	8.0 (5.0; 12.0)	10.0 (7.0; 15.0)	10.0 (5.0; 21.0)
TJC68	11.0 (7.0; 16.0)	14.0 (8.0; 21.0)	18.0 (11.0; 27.0)
Physician DA VAS, mm	59.0 (46.0; 70.0)	65.0 (52.0; 76.0)	63.0 (52.0; 78.0)
Patient pain VAS, mm	63.0 (47.0; 77.0)	69.0 (57.0; 82.0)	68.0 (52.0; 81.0)
HAQ-DI	1.13 (0.69; 1.63)	1.38 (0.75; 1.88)	1.13 (0.63; 1.63)

Data are presented as median (IQR) unless otherwise stated. *All patients were biological DMARD naive as defined by the study protocol. †Medication received at or begun after the start of study treatment and before the last regular visit. ‡Relative frequencies based on the number of patients with an available RF or anti-CCP value.

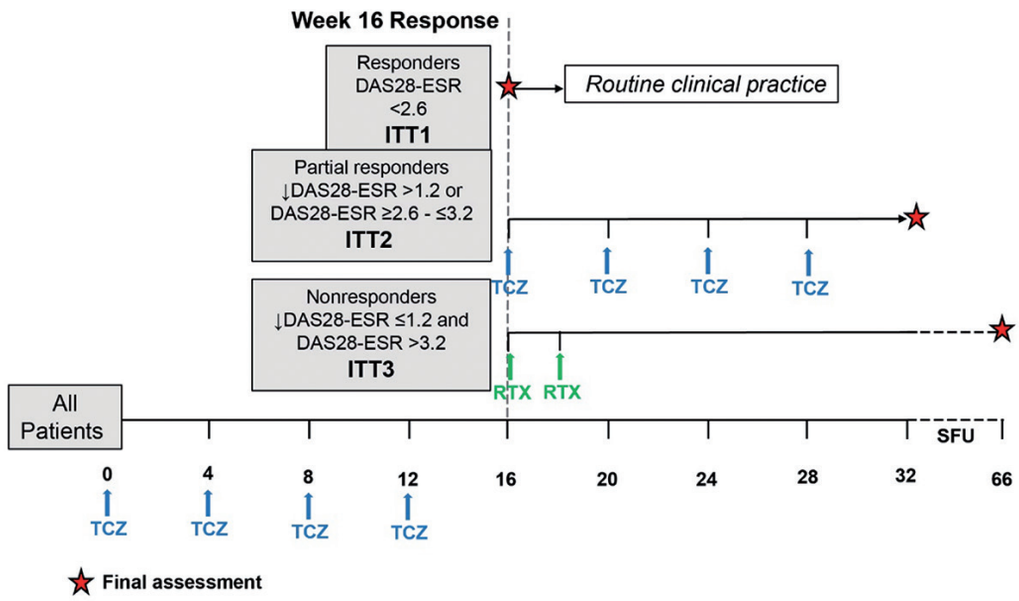
BMI: body mass index; CCP: cyclic citrullinated peptide; csDMARD: conventional synthetic DMARD; DA: disease activity; HAQ-DI: Health Assessment Questionnaire–Disability Index; IQR: interquartile range; ITT1: patients who received ≥1 dose of TCZ in the first treatment period and completed the study, achieving DAS28-ESR remission, at week 16; ITT2: patients who received ≥1 dose of TCZ in the first treatment period with ≥1 efficacy measurement under TCZ, receiving TCZ in the second treatment period; ITT3: patients who received ≥1 dose of TCZ in the first period, ≥1 dose of RTX in the second period, and had ≥1 efficacy measurement during RTX treatment; RF: rheumatoid factor; RTX: rituximab; SD: standard deviation; SJC66: swollen joint count based on 66 joints; TCZ: tocilizumab; TJC68: tender joint count based on 68 joints; VAS: visual analogue scale.

Supplementary Table SII. Efficacy in the subpopulations.

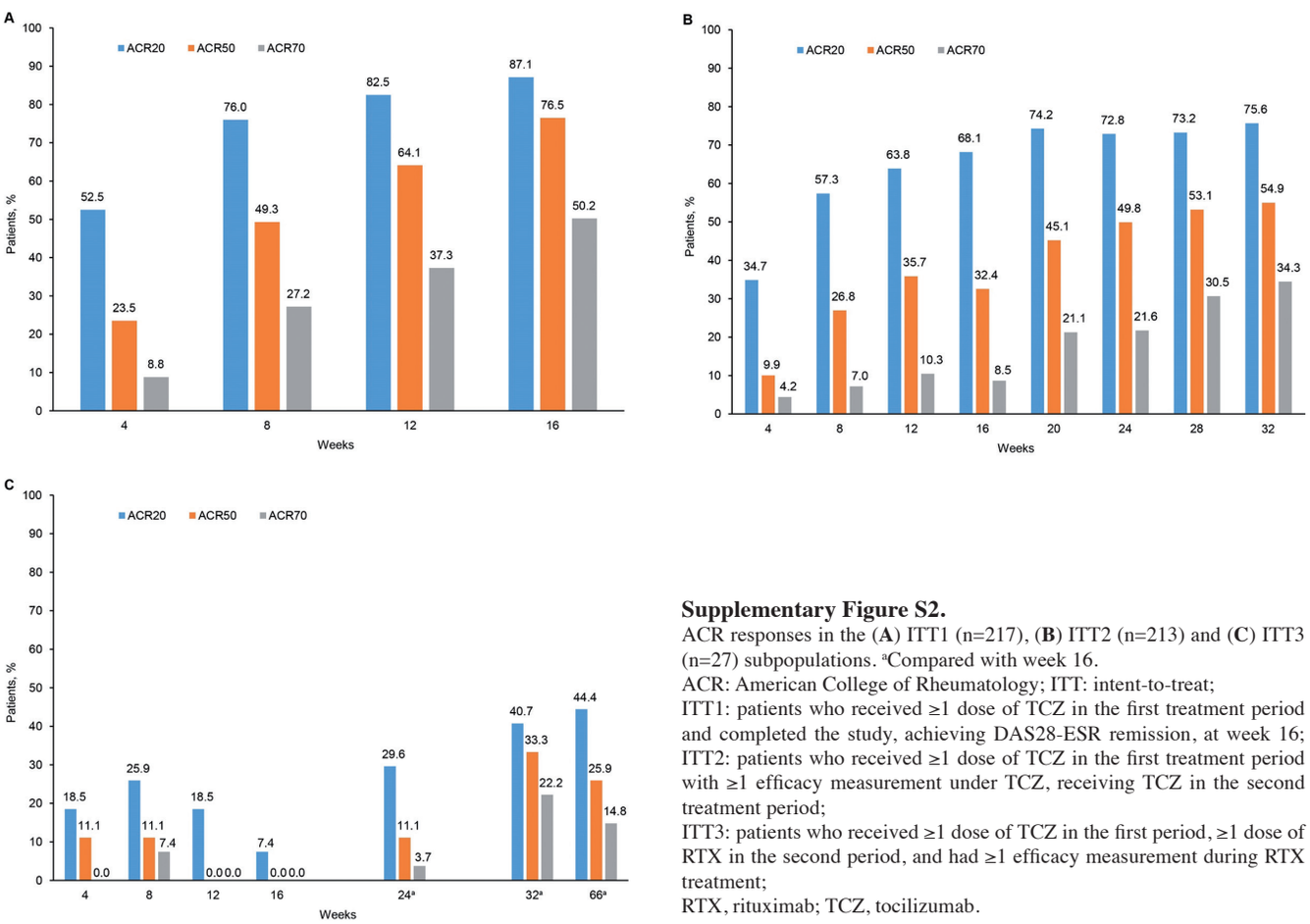
Variable	ITT1[§] n=217		ITT2[§] n=213		ITT3[§] n=27	
DAS28-ESR <2.6, n (%)						
Week 8	145 (66.8)		49 (23.0)		4 (14.8)	
Week 16	217 (100.0)		3 (1.4)		0 (0.0)	
Week 32	NA		117 (54.9)		4 (14.8)	
Week 66*	NA		NA		6 (22.2)	
RF status[†]	RF^+ n=137	RF^- n=61	RF^+ n=144	RF^- n=54	$RF^{+¶}$ n=17	$RF^{-¶}$ n=9
Week 16	137 (100)	61 (100)	3 (2.1)	0 (0.0)	–	–
Week 32	NA	NA	84 (58.3)	25 (46.3)	–	–
Week 66*	NA	NA	NA	NA	–	–
aCCP status[†]	$aCCP^+$ n=110	$aCCP^-$ n=34	$aCCP^+$ n=102	$aCCP^-$ n=35	$aCCP^{+¶}$ n=14	$aCCP^{-¶}$ n=4
Week 16	110 (100)	34 (100)	3 (2.9)	0 (0.0)	–	–
Week 32	NA	NA	60 (58.8)	13 (37.1)	–	–
Week 66*	NA	NA	NA	NA	–	–
DAS28-ESR ≤3.2, n (%)						
Week 16	217 (100.0)		133 (62.4)		0 (0.0)	
Week 32	NA		142 (66.7)		9 (33.3)	
Week 66*	NA		NA		8 (29.6)	
DAS28-ESR, median (IQR)						
Baseline	5.3 (4.6; 6.1)		5.9 (5.4; 6.6)		5.7 (4.8; 6.4)	
Change to week 16[‡]	–3.8 (–4.8; –2.9)		–2.7 (–3.3; –2.0)		–0.7 (–1.1; –0.1)	
Change to week 32[‡]	NA		–3.6 (–4.5; –2.7)		–1.7(–2.6; –0.6)	
Change to week 66*,[‡]	NA		NA		–1.3 (–2.9; –0.5)	
CDAI ≤10.0, n (%)						
Week 16	196 (90.3)		58 (27.2)		0 (0.0)	
Week 32	NA		126 (59.2)		12 (44.4)	
Week 66*	NA		NA		9 (33.3)	
CDAI ≤2.8, n (%)						
Week 16	85 (39.2)		0 (0.0)		0 (0.0)	
Week 32	NA		41 (19.2)		4 (14.8)	
Week 66*	NA		NA		5 (18.5)	
CDAI, median (IQR)						
Baseline	25.7 (20.3; 32.4)		31.9 (25.8; 40.4)		33.5 (21.8; 42.8)	
Change to week 16[‡]	–21.5 (–29.3; –15.1)		–16.9 (–25.4; –11.2)		–6.3 (–9.8; 2.0)	
Change to week 32[‡]	NA		–22.8 (–32.0; –16.3)		–15.2 (–28.7; –8.2)	
Change to week 66*,[‡]	NA		NA		–13.7 (–30.4; –9.1)	
SDAI, median (IQR)						
Baseline	27.8 (21.4; 34.7)		33.3 (27.0; 43.2)		34.5 (21.9; 43.1)	
Change to week 16[‡]	–23.2 (–31.4; –16.3)		–18.1 (–27.7; –12.7)		–6.7 (–9.8; 0.4)	
Change to week 32[‡]	NA		–23.7 (–32.7; –17.2)		–17.3 (–30.9; –7.7)	
Change to week 66*,[‡]	NA		NA		–15.5 (–31.6; –8.2)	
ACR/EULAR 2011 Boolean remission, n (%)						
Week 16	72 (33.2)		2 (0.9)		0 (0.0)	
Week 32	NA		40 (18.8)		6 (22.2)	
Week 66*	NA		NA		3 (11.1)	
CRP, mg/L, mean (SD)						
Baseline	17 (18)		16 (18)		16 (20)	
Week 16	2 (2)		3 (7)		8 (14)	
Week 32	NA		2 (5)		12 (19)	
Week 66*	NA		NA		13 (20)	

Despite being in remission at week 16, three patients in the ITT2 group erroneously received a TCZ infusion. *Observational safety follow-up until week 66 was performed in 27 patients who received subsequent RTX. [†]Based on the number of patients with an available RF or anti-CCP value. [‡]Compared with baseline. [§]ITT1, ITT2 and ITT3 are subpopulations of the overall study population (N=519). [¶]Because of the small sample size of the ITT3 subpopulation, robust comparisons of DAS28 remission by RF and aCCP status were not possible.

ACR: American College of Rheumatology; CDAI: Clinical Disease Activity Index; ITT1: patients who received ≥1 dose of TCZ in the first treatment period and completed the study, achieving DAS28-ESR remission, at week 16; ITT2: patients who received ≥1 dose of TCZ in the first treatment period with ≥1 efficacy measurement under TCZ, receiving TCZ in the second treatment period; ITT3: patients who received ≥1 dose of TCZ in the first period, ≥1 dose of RTX in the second period, and had ≥1 efficacy measurement during RTX treatment; NA: not applicable; RF: rheumatoid factor; RTX: rituximab; SDAI: Simplified Disease Activity Index; TCZ: tocilizumab.



Supplementary Figure S1.
Patient subpopulation treatment algorithm. TCZ was administered intravenously at 8 mg/kg; RTX was administered intravenously at 1000 mg. The SFU was an observational non-treatment period during which patients treated with RTX received treatment according to best medical judgement.
DAS28-ESR: Disease Activity Score based on 28 joints and erythrocyte sedimentation rate; ITT: intent-to-treat; RTX: rituximab; SFU: safety follow-up; TCZ: tocilizumab.



Supplementary Figure S2.
ACR responses in the (**A**) ITT1 (n=217), (**B**) ITT2 (n=213) and (**C**) ITT3 (n=27) subpopulations. [a]Compared with week 16.
ACR: American College of Rheumatology; ITT: intent-to-treat;
ITT1: patients who received ≥1 dose of TCZ in the first treatment period and completed the study, achieving DAS28-ESR remission, at week 16;
ITT2: patients who received ≥1 dose of TCZ in the first treatment period with ≥1 efficacy measurement under TCZ, receiving TCZ in the second treatment period;
ITT3: patients who received ≥1 dose of TCZ in the first period, ≥1 dose of RTX in the second period, and had ≥1 efficacy measurement during RTX treatment;
RTX, rituximab; TCZ, tocilizumab.