RMD Open

Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Tofacitinib versus tocilizumab in the treatment of biological-naïve or previous biological-failure patients with methotrexate-refractory active rheumatoid arthritis

Shunsuke Mori [iD],[1] Yukitomo Urata [iD],[2] Tamami Yoshitama,[3] Yukitaka Ueki[4]

Correspondence to
Dr Shunsuke Mori;
mori.shunsuke.ra@mail.hosp.go.jp

ABSTRACT

Objectives To compare effectiveness between tofacitinib and tocilizumab treatments for biological disease-modifying antirheumatic drug (bDMARD)-naïve patients or previous bDMARD-failure patients with active rheumatoid arthritis (RA) refractory to methotrexate (MTX).

Methods We used two ongoing real-world registries of patients with RA who had first started tofacitinib or tocilizumab between August 2013 and February 2019 at our institutions. Clinical disease activity index (CDAI)-based improvements at 12 months were used for comparisons between tofacitinib and tocilizumab treatments, separately for bDMARD-naïve and previous bDMARD-failure patients.

Results A total of 464 patients with RA with high or moderate CDAI were enrolled (247 with tofacitinib and 217 with tocilizumab). After adjustments for treatment-selection bias by propensity score matching, we showed that tofacitinib was more likely to induce and maintain ≥85% improvement in CDAI (CDAI85), CDAI70 and remission at 12 months compared with tocilizumab in bDMARD-naïve patients. After adjusting for concurrent use of MTX and prednisolone, the ORs of tofacitinib versus tocilizumab were 3.88 (95% CI 1.87 to 8.03) for CDAI85, 2.89 (95% CI 1.43 to 5.84) for CDAI70 and 3.31 (95% CI 1.69 to 6.48) for remission. These effects were not observed in bDMARD-failure patients. In tofacitinib treatment for bDMARD-failure patients, the number of previously failed bDMARD classes was not associated with CDAI-based improvements. The rate of overall adverse events was similar between both treatments. Similar ORs were obtained from patients adjusted by inverse probability of treatment weighting.

Conclusions Compared with tocilizumab, tofacitinib can induce greater improvements during the first 12-month treatment in bDMARD-naïve patients, but this difference was not observed in previous bDMARD-failure patients.

INTRODUCTION

Tofacitinib, a potent selective inhibitor of Janus kinases (JAKs), is the first targeted synthetic disease-modifying antirheumatic drug (DMARD) approved for treatment of

Key messages

What is already known about this subject?

► To determine the optimal position of tofacitinib in the treatment algorithm for rheumatoid arthritis (RA), we need to compare its initial efficacy and safety with those of biological disease-modifying antirheumatic drugs (bDMARDs). In recent clinical trials, tofacitinib had at least equal efficacy and similar safety to adalimumab, an antitumour necrosis factor antibody, in the treatment of patients with rheumatoid arthritis (RA) with inadequate response to methotrexate (MTX). Currently, there are few studies regarding the comparison of effectiveness and safety between tofacitinib and the anti-interleukin 6 receptor antibody tocilizumab.

What does this study add?

► In this multicentre cohort study using two ongoing real-world registries, tofacitinib was more likely to induce and maintain ≥85% improvement in clinical disease activity index (CDAI85), CDAI70 and remission during the first 12-month treatment compared with tocilizumab in bDMARD-naïve patients with active RA despite MTX therapy; these differences were not observed in the treatment of previous bDMARD-failure patients.

► Among tofacitinib-treated patients, the number of failed bDMARD classes was not associated with CDAI-based improvements.

► There was no significant difference in the rate of overall adverse events that caused drug discontinuation between tofacitinib and tocilizumab treatments or bDMARD-naïve and failure patients.

How might this impact on clinical practice or further developments?

► In current practice, biological therapy is preferentially used in patients with RA who have had an inadequate response to MTX, but considering the greater effectiveness of tofacitinib in bDMARD-naïve patients than tocilizumab, tofacitinib could be considered as another option for such patients with RA prior to the start of biological therapy.

rheumatoid arthritis (RA).[1–3] Recent phase III clinical trials showed that tofacitinib is effective and generally well tolerated in the treatment of active RA, both as monotherapy and in combination with methotrexate (MTX) or other conventional synthetic DMARDs (csDMARDs).[4–11] Long-term extension studies demonstrated stable safety and sustained efficacy.[12–16]

For optimising RA management, we need to clarify whether tofacitinib should be considered as an option only for patients who have failed to respond to at least one biological DMARD (bDMARD) or for MTX-resistant or MTX-intolerant patients prior to any attempted biological therapy. In a previous study, we found that the effect of tofacitinib on initial improvement is significantly higher in bDMARD-naïve patients than in patients who failed previous bDMARD therapy.[17] However, the number of patients included in that study was small and the effect of previous bDMARD number on outcomes of tofacitinib therapy was unclear. Recent phase III and IIIb/IV trials revealed that tofacitinib has at least equal efficacy and similar safety to adalimumab, an anti-tumour necrosis factor (TNF)-α antibody, in the treatment of patients with RA with inadequate response to MTX.[5 10] There are few comparison studies between tofacitinib and non-TNF inhibitors such as tocilizumab, an anti-interleukin 6 (IL-6) receptor antibody, in regard to their effectiveness and safety, however.

In the present study, we used ongoing real-world registries including patients with MTX-refractory active RA who had first begun tofacitinib or tocilizumab in our institutions between August 2013 and February 2019. We compared therapeutic outcomes at 12 months between tofacitinib-treated and tocilizumab-treated patients, grouped by status of previous bDMARD use, after propensity score (PS) matching.

METHODS
Patients
This was a multicentre cohort study in collaboration with the following institutions in Japan: National Hospital Organization Kumamoto Saishun Medical Center, Tsugaru General Hospital United Municipalities of Tsugaru, Sasebo Chuo Hospital and Yoshitama Clinic for Rheumatic Diseases. This cohort was a group that consisted of patients with RA who had first started tofacitinib or tocilizumab between August 2013 and February 2019 at the rheumatology unit of these institutions. All new users of tofacitinib and tocilizumab were prospectively enrolled in ongoing multicentre registries, namely, the TOFARA registry (the TOFacitinib treatment for Active RA registry) and the ACTRA-RI registry (ACTemura for RA patients with or without Renal Insufficiency), respectively.[18 19] Data at baseline and during follow-up were regularly deposited in the databases of these registries.

Participants in this study were required to be over 18 years of age at the time of tofacitinib or tocilizumab initiation, to fulfil the 1987 American College of Rheumatology (ACR) criteria or the 2010 ACR/EULAR) criteria for diagnosis of RA,[20 21] to have shown an insufficient improvement (less than 50% of initial disease activity) despite MTX therapy for ≥3 months (defined as MTX-refractory RA), and to have had a high or moderate clinical disease activity index (CDAI>10) at the start of treatment with tofacitinib or tocilizumab. In addition, patients who had discontinued bDMARD therapy due to intolerance were excluded from this study.

Study design
Tofacitinib was administered at a dosage of 5 mg two times per day or 5 mg once a day as per the licensed dose for RA in Japan, and tocilizumab was administered by intravenous infusion at 8 mg/kg every 4 weeks or by a subcutaneous injection of 162 mg every other week. For patients who failed to respond to a bDMARD, the start of tofacitinib or tocilizumab treatment was delayed at least until the next-scheduled dosing day for that bDMARD in accordance with the manufacturer's instructions. All participants in this study were MTX refractory. After the treatment selection of tofacitinib or tocilizumab, we recommended continuing or restarting MTX as a background therapy during both treatments unless patients refused it. The treatment selection of tofacitinib or tocilizumab was not influenced by the concurrent use of MTX, because it was recognised that, unlike TNF-α inhibitors, tofacitinib and tocilizumab monotherapies generally have good clinical efficacy.[22] For patients who chose concurrent use of MTX at the start of tofacitinib or tocilizumab treatment, we principally continued it at stable doses (4–14 mg/week) during follow-up. Concurrent use of prednisolone (PSL, 2–7.5 mg/day) during tofacitinib or tocilizumab therapy was left to the treating physicians' discretion. We avoided prolonged use of PSL and used it for priming when initiating tofacitinib or tocilizumab, regardless of tofacitinib or tocilizumab therapy. If a patient failed to have controlled disease activity despite the initial regimen, including the status of MTX use, we categorised that patient as a dropout due to lack or loss of efficacy, and follow-up was ended. As a result, most patients adhered to the initial status of MTX use during follow-up.

Patients assigned to the tofacitinib and tocilizumab groups were further divided into two patient groups according to status of previous bDMARD use, namely, bDMARD-naïve patients and previous bDMARD-failure patients. Previous bDMARD-failure patients were defined as patients who had experienced an inadequate response (lack or loss of efficacy) to at least one bDMARD. Previous bDMARD-failure patients were further classified into patients with one-class failure and two-class or three-class failure. The classes of bDMARDs included the following three groups: TNF-α inhibitors (infliximab, etanercept, adalimumab, golimumab and certolizumab pegol), an IL-6 inhibitor (tocilizumab) and a T-cell signalling inhibitor (abatacept).

Baseline characteristics, including age, gender, RA duration, Steinbrocker radiological stage, anticyclic citrullinated peptide antibodies (anti-CCP)-positive status, rheumatoid factor (RF)-positive status and CDAI values, were recorded for all participants at registry enrolment.

Therapeutic outcomes were examined every 4 weeks. The end of follow-up was set as the time of dropout from the study or 12 months after initiation of the tofacitinib or tocilizumab therapy, whichever came first. Dropout from the study was defined as drug discontinuation because of an adverse event, lack or loss of efficacy, or lost to follow-up before the end of the 12-month treatment. Patients who had missed at least two scheduled visits would have been classified as lost to follow-up (patient preference, hospital transfer, surgery, etc). Therapeutic outcomes at month 12 were compared between tofacitinib-treated and tocilizumab-treated patients after stratification according to previous use of bDMARDs.

CDAI-based improvement measures
The CDAI was used to quantify RA disease activity during treatment with tofacitinib or tocilizumab because the CDAI does not include acute-phase reactant values.[17 23] Cut-off values for disease activity states were defined as follows: high disease activity, CDAI>22; moderate disease activity, CDAI>10 and ≤22; low disease activity, CDAI>2.8 and ≤10 and remission, CDAI≤2.8.[24] Therapeutic response was evaluated according to the CDAI improvement criteria, in which minor, moderate and major response are defined as ≥50% (CDAI50), ≥70% (CDAI70) and ≥85% (CDAI85) improvements in CDAI, respectively.[25] CDAI improvement based on the minimum clinically important difference (MCID), defined as a CDAI reduction >12 for patients starting with a high CDAI and >6 for those starting with a moderate CDAI, was also assessed.[26]

Safety analysis
Adverse events were classified according to the system organ classes described in the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0.

Patient and public involvement
Patients were not involved in the design or conduct of the study, development of outcomes or dissemination of study results.

Statistical analysis
The baseline characteristics and CDAI-based improvement measures at 12 months were compared using the independent-measures t-test for continuous variables and the χ^2 test for categorical variables. After PS matching, baseline characteristics and CDAI-based improvements at 12 months were compared using the paired-sample t-test for continuous variables and the McNemar test for categorical variables. If data deviated substantially from the normal distribution, we used bootstrapping for comparisons of continuous variables.

The primary outcome was the proportion of patients who had achieved and maintained each of the CDAI-based improvements at 12 months. To classify patients according to CDAI-based improvements at 12 months (CDAI85, CDAI70, CDAI50 and MCID-based improvement), non-responder imputation (NRI) was used for dropout patients. The rates of dropout patients due to adverse events, lack or loss of efficacy, or lost to follow-up in each of the treatment groups was also compared. To calculate mean CDAI values at 12 months, missing data on dropout patients were imputed using baseline observation carried forward (BOCF). Multiple imputation (MI) was conducted to examine the impact of NRI on the robustness of results. Twenty imputed datasets were generated based on logistic regression models for binary endpoints (each of the CDAI-based improvements) or linear regression models for continuous endpoints (CDAI values). The multiply imputed datasets were analysed using the same statistical method as those used for the primary analysis. The results from each of imputed datasets were combined using Rubin's rule. All available data at postbaseline visits up to the end of follow-up were included in the model.

PS matching was used to adjust for treatment-selection bias. The PS was the probability of a patient receiving the treatment being tested (tofacitinib) based on observed baseline covariates, which was calculated for each patient using a logistic regression model. Each patient treated with tofacitinib was allocated to one tocilizumab-treated patient (1:1 matching), with the same PS or with a PS that differed only slightly, using a calliper width equal to 0.02 of the SD of the logit of the PS.[27 28] The matching algorithm was the nearest neighbour matching without replacement. The balance of baseline characteristics between the two treatment groups after PS matching was assessed using the absolute standardised difference (ASD) of each covariate, and an ASD<0.10 was considered a negligible difference in the mean or prevalence of a covariate between both groups.[29]

Adjusted ORs and 95% CIs for treatment effect (tofacitinib vs tocilizumab) on each of the CDAI-based improvements were calculated using conditional multivariable logistic regression analyses. For calculating adjusted ORs and 95% CIs for the effect of previous failures to bDMARDs on the CDAI-based improvements in tofacitinib-treated patients, standard multivariable logistic regression analyses were used.

As sensitive analyses, we conducted inverse provability of treatment weighting (IPTW) based on PS to receive the treatment (tofacitinib), in which each patient was weighted by the reciprocal of the probability of receiving the treatment (1/PS for tofacitinib-treated and 1/[1−PS] for tocilizumab-treated subject). We calculated ASD to compare the balance in baseline covariates between the treatment groups, and the threshold was set to be <0.10.[30 31] The IPTW method enabled the retention of all patients in the analysis.[28 31] After the IPTW adjustment, ORs and 95% CIs for treatment effect (tofacitinib vs

tocilizumab) on each of the CDAI-based improvements were calculated using univariable logistic regression analysis.

To adjust for possible time-varying covariates (concurrent use of MTX and PSL) during the 1-year follow-up, we used the IPTW method, in which previous exposure history was incorporated in the PS estimation at each time point. Age, gender, RA duration, radiological stage, anti-CCP status, RF status, initial CDAI values, MTX use and PSL use were determined at baseline. MTX use and PSL use were measured at 3, 6, 9 and 11 months. Missing data on dropout patients at each timepoint were imputed using MI. First, we imputed CDAI values at each timepoint based on linear regression models. Next, imputed data of MTX and PSL use were determined with logistic regression models. Fifty-one imputed datasets were generated. IPTW weights were calculated based on the probability of exposure at each timepoint conditioned on the exposure history of the previous timepoint, the time-varying covariates history at the current timepoint and the non-time-varying covariates. A single IPTW for each of the imputed datasets was generated based on these weights. To reduce variability due to instability in estimation that can be induced by subjects with very large weights, we used stabilised IPTW weights, which were calculated by multiplying IPTW weights by the estimated probability of exposure at each timepoint conditioned on the exposure

history of the previous timepoint and non-time-varying covariates. The non-time-varying covariates included age, gender, RA duration, radiological stage, anti-CCP status, RF status and initial CDAI values. The final stabilised IPTW was used as the weight in logistic regression modelling with the outcome at 12 months. The ORs and 95% CIs for treatment effect (tofacitinib vs tocilizumab) on each of the CDAI-based improvements were calculated. The results from different imputed datasets were pooled to form a set of single estimates using Rubin's formula. We calculated ASD to compare the balance in the baseline covariates and the MTX and PSL use at each timepoint between treatment groups.[32–36]

All calculations were performed using PASW Statistics V.22 (SPSS Japan), STATA release V.16 (StataCorp) and Easy R (Saitama Medical Center, Jichi Medical University, Saitama, Japan).[37]

RESULTS

Patient characteristics at baseline

A total of 464 patients with MTX-refractory RA with high or moderate CDAI who had started treatment with tofacitinib (247 patients) or tocilizumab (217 patients) for the first time in participating institutions between August 2013 and February 2019 were included in the present study (figure 1). Among participants, 215 were
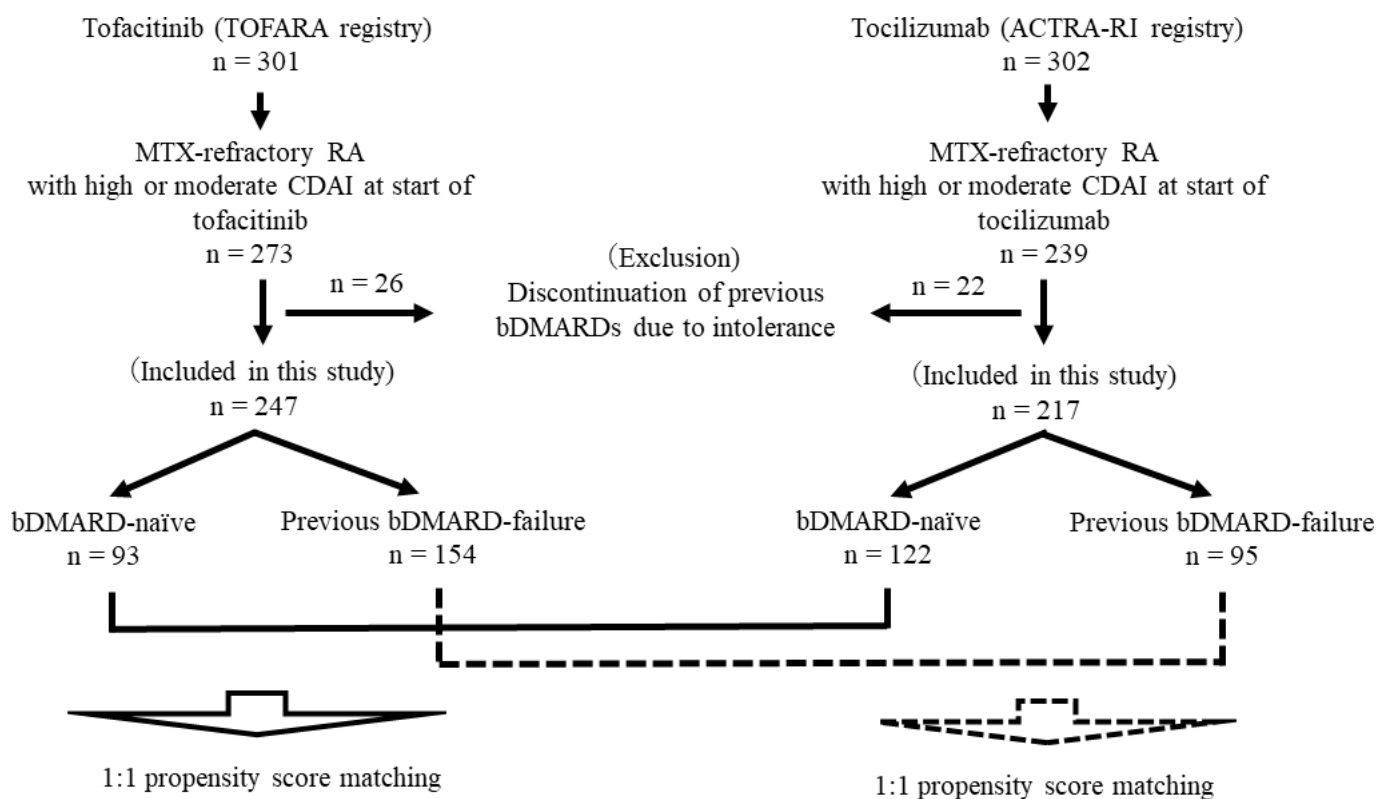

Figure 1 Flowchart of study. MTX-refractory patients were defined as patients who had shown insufficient improvement (less than 50% of CDAI) to MTX therapy for ≥3 months. Previous bDMARD-failure patients were defined as patients who had experienced an inadequate response (lack or loss of efficacy) to at least one bDMARD. ACTRA-RI, ACTemura for RA patients with or without Renal Insufficiency; bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MTX, methotrexate; RA, rheumatoid arthritis; TOFARA, TOFacitinib treatment for Active RA.

Table 1 Baseline characteristics of tofacitinib-treated and tocilizumab-treated patients with RA, grouped by the status of previous bDMARD use, before propensity score matching

| | bDMARD-naïve patients (n=215) | | | | Previous bDMARD-failure patients (n=249) | | | |
	Tofacitinib (n=93)	Tocilizumab (n=122)	P value*	ASD†	Tofacitinib (n=154)	Tocilizumab (n=95)	P value*	ASD†
Baseline characteristics								
Age, years, mean (SD)	63.6 (13.7)	64.8 (13.6)	0.52	0.088	66.7 (11.5)	61.6 (12.5)	0.001	0.43
Male sex, number (%)	20 (21.5)	39 (32.0)	0.093	0.20	33 (21.4)	18 (18.9)	0.75	0.051
RA duration, years, mean (SD)	8.8 (10.4)	7.8 (8.9)	0.47	0.10	13.9 (9.7)	11.2 (8.8)	0.030	0.29
Advanced stage‡, number (%)	31 (33.3)	33 (27.0)	0.37	0.11	97 (63.0)	50 (52.6)	0.11	0.17
Anti-CCP-positive, number (%)	78 (83.9)	97 (79.5)	0.48	0.095	135 (87.7)	83 (87.4)	1.00	0.007
RF-positive, number (%)	81 (87.1)	99 (81.1)	0.27	0.15	128 (83.1)	81 (85.3)	0.72	0.049
CDAI, mean (SD)	22.7 (9.6)	23.0 (10.4)	0.82	0.030	23.1 (10.2)	22.6 (9.9)	0.73	0.050
High CDAI (>22), number (%)	38 (40.9)	55 (45.1)	0.58	0.070	76 (49.4)	39 (41.1)	0.24	0.14
Concurrent MTX use, number (%)	63 (67.7)	75 (61.5)	0.39	0.11	84 (54.5)	66 (69.5)	0.023	0.25
Concurrent PSL use, number (%)	27 (29.0)	64 (52.5)	0.001	0.41	42 (27.3)	44 (46.3)	0.003	0.33

*Comparisons of each of the baseline characteristics between the tofacitinib and tocilizumab groups, separately in bDMARD-naïve patients and bDMARD-failure patients, using the independent-measures t-test for continuous variables and the χ^2 test for categorical variables.
†ASD of <0.1 indicates that the baseline characteristic was well balanced between the two treatment groups (tofacitinib vs tocilizumab).
‡Advanced stage was defined as Steinbrocker radiological stages III and IV.
anti-CCP, anti-cyclic citrullinated peptide antibodies; ASD, absolute standardised difference; bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MTX, methotrexate; PSL, prednisolone; RA, rheumatoid arthritis; RF, rheumatoid factor.

bDMARD-naïve patients (93 in the tofacitinib group and 122 in the tocilizumab group), and 249 had experienced an inadequate response to one or more previous bDMARDs (154 in the tofacitinib group and 95 in the tocilizumab group). All bDMARD-failure patients started treatment with tofacitinib or tofacitinib after the scheduled dosing intervals of currently used bDMARDs had passed according to the manufacturer's instructions. In the present study, none of the patients who started tocilizumab had previously received tofacitinib therapy, and 30 patients who started tofacitinib had experienced failure with tocilizumab therapy. The baseline characteristics of all participants are presented in table 1.

To adjust for treatment-selection bias, we conducted 1:1 PS matching of baseline characteristics, including age, gender, RA duration, radiological stage, anti-CCP status, RF status and initial CDAI values, between the tofacitinib and tocilizumab groups, separately for bDMARD-naïve patients and bDMARD-failure patients. The concurrent use of MTX or PSL was not matched with the PS because the use of these drugs was determined after treatment selection and did not influence the selection. According to ASD values, the balance of PS-matched baseline characteristics between the treatment groups was improved (table 2).

Comparisons of CDAI-based improvements at 12 months between the tofacitinib and tocilizumab groups

For bDMARD-naïve patients after PS matching, tofacitinib-treated patients were more likely to achieve and maintain remission during 12 months than tofacitinib-treated

patients (table 3). The rates of CDAI85 and CDAI70 responses were significantly higher in the tofacitinib group compared with the tocilizumab group. The rates of CDAI50 and MCID-based CDAI improvement at 12 months were similar between the treatment groups. For bDMARD-failure patients, the rate of patients achieving and maintaining low CDAI during 12 months was significantly higher in the tofacitinib group, but there were no significant differences in the rates of patients who achieved and maintained remission, CDAI85, CDAI70, CDAI500 or MCID-based CDAI improvement at 12 months between both treatment groups.

As shown in table 3, the rates of dropout due to lost to follow-up were approximately 1% of all patients in each treatment group. The rates of drop-out patients due to adverse events ranged from 5% to 7.5% of all patients in each treatment group. Among them, half of the patients failed to have controlled disease activity at the time of dropout despite tofacitinib or tocilizumab treatment for 3 months or more. As for dropout patients due to lack or loss of efficacy, their disease activity remained at high/moderate levels or flared up despite treatment for 3 months or more. Between the tofacitinib and tocilizumab treatment groups, there were no significant differences in rates of patients who had discontinued the treatment due to lost to follow-up, lack/loss of efficacy or adverse events.

Unadjusted and adjusted ORs of treatment effect (tofacitinib vs tocilizumab) for each of the CDAI-based improvements were calculated using conditional

Table 2 Baseline characteristics of tofacitinib-treated and tocilizumab-treated patients with RA, grouped by the status of previous bDMARD use, after propensity score matching

	bDMARD-naïve patients (n=186)				Previous bDMARD-failure patients (n=160)			
	Tofacitinib (n=93)	Tocilizumab (n=93)	P value*	ASD†	Tofacitinib (n=80)	Tocilizumab (n=80)	P value*	ASD†
Baseline characteristics								
Age, years, mean (SD)	63.6 (13.7)	64.5 (14.1)	0.53	0.065	65.4 (11.3)	64.7 (10.0)	0.54	0.066
Male sex, number (%)	20 (21.5)	21 (22.6)	1.00	0.022	18 (22.5)	17 (21.3)	1.00	0.024
RA duration, years, mean (SD)	8.8 (10.4)	8.3 (9.2)	0.70	0.051	11.9 (8.5)	12.1 (9.0)	0.84	0.023
Advanced stages‡, number (%)	31 (33.3)	31 (33.3)	1.00	<0.001	46 (57.5)	46 (57.5)	1.00	<0.001
Anti-CCP-positive, number (%)	78 (83.9)	76 (81.7)	0.84	0.047	72 (90.0)	69 (86.3)	0.61	0.092
RF-positive, number (%)	81 (87.1)	79 (84.9)	0.83	0.051	67 (83.8)	67 (83.8)	1.00	<0.001
CDAI, mean (SD)	22.7 (9.6)	22.9 (10.6)	0.89	0.020	22.5 (10.6)	23.2 (9.7)	0.69	0.020
High CDAI (>22), number (%)	38 (40.9)	41 (44.1)	0.74	0.053	36 (45.0)	36 (45.0)	1.00	<0.001
Concurrent MTX use§, number (%)	63 (67.7)	53 (57.0)	0.14	0.18§	40 (50)	54 (67.5)	0.04	0.30§
Concurrent PSL use§, number (%)	27 (29.0)	48 (51.6)	0.005	0.38§	28 (35.0)	38 (47.5)	0.14	0.21§

*Comparisons of each of the baseline characteristics between the tofacitinib and tocilizumab groups after propensity score matching, separately in bDMARD-naïve patients and bDMARD-failure patients, using the paired-sample t-test for continuous variables and the McNemar test for categorical variables.

†ASD of <0.10 indicates that the baseline characteristic was well balanced between the two treatment groups (tofacitinib vs tocilizumab).

‡Advanced stages were defined as Steinbrocker radiological stages III and IV.

§Concurrent use of MTX and PSL was not used in propensity score matching because the use of these drugs was determined after the treatment selection.

anti-CCP, anti-cyclic citrullinated peptide antibodies; ASD, absolute standardised difference; bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MTX, methotrexate; PSL, prednisolone; RA, rheumatoid arthritis; RF, rheumatoid factor.

univariable and multivariable logistic regression analyses (table 4). For bDMARD-naïve patients, the final logistic regression models showed that ORs (95% CIs) of tofacitinib versus tocilizumab adjusted for concurrent MTX and PSL use were 3.31 (1.69 to 6.48) for remission (p<0.001), 3.88 (1.87 to 8.03) for CDAI85 (p<0.001) and 2.89 (1.43 to 5.84) for CDAI70 (p=0.003). For bDMARD-failure patients, after adjustments for concurrent MTX and PSL use, tofacitinib treatment did not induce more favourable effects on any CDAI-based improvements at 12 months compared with tocilizumab treatment. Unadjusted and adjusted ORs of concurrent MTX and PSL use for each of the CDAI-based improvements are shown in online supplemental tables 1 and 2.

To verify the robustness of the primary results, we performed sensitivity analyses using the MI procedure. As shown in supplementary tables (online supplemental tables 3 and 4), the results of these analyses were consistent with those of the primary analyses.

Comparisons of CDAI-based improvements at 12-month tofacitinib treatment between patients with one-class bDMARD failure and those with two-class or three-class bDMARD failure

In the present study, the tofacitinib group with previous bDMARD failure included 154 patients who had experienced an inadequate response to one or more classes of bDMARDs (78 patients with one-class bDMARD failure (50.6%) and 76 with two-class or three-class bDMARD failure (49.4%)). The tocilizumab group with previous bDMARD failure included 95 bDMARD-failure patients

(94 patients with one-class failure (98.9%) and one patient with two-class failure (1.1%)). Compared with the tocilizumab group, the tofacitinib group contained two-class or three-class failure patients at a significantly higher rate. Accordingly, it was difficult to match the number of previously failed bDMARD classes between the tofacitinib and the tocilizumab groups. We needed to consider the possibility that the higher rate of patients with two-class or three-class bDMARD failure in the tofacitinib group might have negative impacts on therapeutic outcomes at 12 months. To address this concern, we compared CDAI-based improvements between tofacitinib-treated patients with one-class bDMARD failure and those with two-class or three-class bDMARD failure.

As shown in table 5, all baseline characteristics except for anti-CCP-positive rates and mean CDAI values were similar among the one-class and two-class or three-class failure groups. There was no significant difference in the rate of any of the CDAI-based improvements between both bDMARD-failure groups. After adjustment for baseline CDAI values, the number of failed bDMARD classes had no negative impacts on the CDAI-based improvements (online supplemental table 5).

Adverse events

Adverse events that caused discontinuation of tofacitinib or tocilizumab therapy are shown in online supplemental table 6. In the tofacitinib group, 22 patients dropped out before the end of the first year due to adverse events (7 bDMARD-naïve and 15 bDMARD-failure patients). In the tocilizumab group, 12 patients dropped out due to

Table 3 Comparisons of therapeutic response at 12 months between tofacitinib-treated and tocilizumab-treated patients with RA after propensity score matching

	bDMARD-naïve patients (n=186)			Previous bDMARD-failure patients (n=160)		
	Tofacitinib (n=93)	Tocilizumab (n=93)	P value*	Tofacitinib (n=80)	Tocilizumab (n=80)	P value*
Therapeutic outcomes at 12 months						
CDAI, mean (SD)†	6.9 (9.9)	9.9 (10.6)	0.030	13.7 (11.7)	14.8 (9.9)	0.55
Dropout, number (%)	17 (18.3)	20 (21.5)	0.69	22 (27.5)	33 (41.3)	0.13
Lack or loss of efficacy	9 (9.7)	13 (14.0)	0.48	17 (21.3)	26 (32.5)	0.19
Adverse events	7 (7.5)	6 (6.5)	1.00	4 (5.0)	6 (7.5)	0.75
Lost to follow-up	1 (1.1)	1 (1.1)	1.00	1 (1.3)	1 (1.3)	1.00
Remission (CDAI ≤2.8), number (%)	53 (57.0)	25 (26.9)	<0.001	10 (12.5)	12 (15.0)	0.83
Low CDAI (>2.8 and ≤10), number (%)	20 (21.5)	37 (39.8)	0.015	35 (43.8)	18 (22.5)	0.007
Remission or low CDAI (≤10), number (%)	73 (78.5)	62 (66.7)	0.082	45 (56.3)	30 (37.5)	0.046
Moderate or high CDAI (>10), number (%)	3 (3.2)	11 (11.8)	0.039	13 (16.3)	17 (21.3)	0.54
Improvements at 12 months, number (%)‡						
CDAI85§ (major response)	59 (63.4)	28 (30.1)	<0.001	13 (16.3)	17 (21.3)	0.58
CDAI70§ (moderate response)	69 (74.2)	45 (48.4)	0.001	26 (32.5)	25 (31.3)	1.00
CDAI50§ (minor response)	71 (76.3)	62 (66.7)	0.18	39 (48.8)	31 (38.8)	0.29
MCID-based CDAI improvement¶	72 (77.4)	63 (67.7)	0.18	44 (55.0)	32 (40.0)	0.11

*Comparisons of CDAI-based improvement measures at 12 months between tofacitinib and tocilizumab therapy after propensity score matching, separately in bDMARD-naïve patients and bDMARD-failure patients, using bootstrapping for continuous variables (bootstrapped paired sample t-test) and the McNemar test for categorical variables. The same p values were also obtained with Wilcoxon signed-rank tests for continuous variables.

†To calculate mean CDAI values at 12 months, missing data on dropout patients were imputed using baseline observation carried forward.

‡For classification at 12 months, non-responder imputation was used for missing data on patients who had withdrawn from the study (dropout patients) because of lack or loss of efficacy, adverse events and lost to follow-up.

§Defined as achieving and maintaining ≥50% improvement of CDAI (CDAI50), ≥70% (CDAI70) and ≥85% (CDAI85) during the 12-month treatment.

¶Defined as CDAI reduction >12 for patients starting with a high CDAI and CDAI reduction >6 for those starting with a moderate CDAI at 12 months of treatment.

bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MCID, minimal clinically important difference; RA, rheumatoid arthritis.

Table 4 Tofacitinib versus tocilizumab: comparison of CDAI-based improvements at 12 months in patients with RA after propensity score matching

	bDMARD-naïve patients with RA				Previous bDMARD-failure patients with RA			
	Unadjusted OR* (95% CI)	P value	Adjusted OR* (95% CI)	P value	Unadjusted OR* (95% CI)	P value	Adjusted OR* (95% CI)	P value
Remission (CDAI≤2.8)	3.33 (1.75 to 6.35)	<0.001	3.31 (1.69 to 6.48)	<0.001	0.83 (0.36 to 1.93)	0.67	0.77 (0.32 to 1.86)	0.56
Remission or low CDAI (≤10)	2.00 (0.97 to 4.12)	0.061	2.05 (0.93 to 4.52)	0.076	1.88 (1.04 to 3.39)	0.035	1.67 (0.90 to 3.12)	0.11
CDAI85† (major response)	4.10 (2.05 to 8.18)	<0.001	3.88 (1.87 to 8.03)	<0.001	0.77 (0.37 to 1.57)	0.47	0.65 (0.29 to 1.45)	0.29
CDAI70† (moderate response)	3.18 (1.62 to 6.26)	<0.001	2.89 (1.43 to 5.84)	0.003	1.05 (0.56 to 1.97)	0.87	0.96 (0.49 to 1.88)	0.90
CDAI50† (minor response)	1.69 (0.85 to 3.36)	0.13	1.56 (0.76 to 3.21)	0.22	1.44 (0.79 to 2.63)	0.23	1.32 (0.70 to 2.48)	0.39
MCID-based improvement‡	1.69 (0.85 to 3.36)	0.13	1.56 (0.76 to 3.21)	0.22	1.67 (0.93 to 2.99)	0.09	1.46 (0.79 to 2.73)	0.23

*Unadjusted ORs (95% CI) of tofacitinib versus tocilizumab were determined for each of the CDAI-based improvement measures according to single conditional logistic regression analyses. ORs (95% CIs) of tofacitinib versus tocilizumab were adjusted for concurrent MTX use and concurrent PSL use by conditional multivariable logistic regression analysis. Unadjusted and adjusted ORs (95% CIs) of confounder variables (MTX use and PSL use) for each of the CDAI-based improvement measures are shown in online supplemental tables 1 and 2.
†Defined as achieving and maintaining ≥50% improvement of CDAI (CDAI50), ≥70% (CDAI70) and ≥85% (CDAI85) during the 12-month treatment.
‡Defined as CDAI reduction >12 for patients starting with a high CDAI and CDAI reduction >6 for those starting with a moderate CDAI at 12 months of treatment.
bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MCID, minimal clinically important difference; MTX, methotrexate; PSL, prednisolone; RA, rheumatoid arthritis.

adverse events (7 bDMARD-naïve and 5 bDMARD-failure patients). There was no significant difference in the rate of overall adverse events between tofacitinib and tocilizumab treatment for bDMARD-naïve or failure patients, before or after PS matching (online supplemental table 6 and table 3). There were no death cases during the follow-up period.

Sensitivity analyses

To adjust for all baseline confounding factors, regardless of whether they affected the treatment assignment or not, without losing the data of PS-unmatched patients, we used the IPTW method based on PS estimated from baseline covariates, including age, gender, RA duration, radiological stage, anti-CCP status, RF status and initial CDAI values as well as the concurrent use of MTX and PSL, separately for bDMARD-naïve patients and bDMARD-failure patients. According to ASD values after IPTW adjustment, the observed baseline covariates were independent of the treatment assignment (online supplemental table 7). Similar ORs of tofacitinib versus tocilizumab were produced between the IPTW-adjusted and PS-matched models (table 6).

Considering the generalisation of the study, we performed sensitivity analysis that included all patients in our cohort (online supplemental table 8). After the IPTW adjustments for treatment-selection bias, we made comparisons of each of the CDAI-based improvements at 12 months between the tofacitinib and tocilizumab groups, separately for bDMARD-naïve and previous bDMARD-experienced patients. ASD values showed a good balance in baseline covariates between the treatments (online supplemental table 9). Similar ORs of tofacitinib versus tocilizumab were obtained when compared with the estimates obtained from PS-matched or IPTW-adjusted patients with RA who had met the inclusion and exclusion criteria of this study (online supplemental table 10).

To adjust for time-varying co-founders, we created a stabilised IPTW based on PS estimations at baseline, 3, 6, 9 and 11 months, and used it as the weight for the adjustments of time-varying covariates in logistic regression analysis for treatment effect on each of the CDAI-based improvements at 12 months. ASD values showed a good balance in the baseline covariates as well as the use of MTX and PSL at each timepoint between the treatments (online supplemental tables 11 and 12). The ORs of tofacitinib versus tocilizumab after the adjustment for time-varying covariates were similar to those obtained without such adjustment (table 7). The influence of the concurrent use of MTX and PSL during the follow-up period appeared limited.

DISCUSSION

In this registry-based cohort study for MTX-refractory patients with RA with high or moderate CDAI, tofacitinib was more likely to induce and maintain CDAI85, CDAI70

Table 5 Baseline characteristics and therapeutic response to 12-month tofacitinib treatment in previous bDMARD-failure patients

	bDMARD-failure patients (n=154)		
	1-class failure (n=78)	2-class or 3-class failure (n=76)	P value*
Baseline characteristics			
Age, years, mean (SD)	66.4 (11.7)	67.1 (11.4)	0.67
Male sex, number (%)	16 (20.5)	17 (22.4)	0.85
RA duration, years, mean (SD)	14.1 (9.9)	13.8 (9.4)	0.82
Advanced stage†, number (%)	48 (61.5)	49 (64.5)	0.74
Anti-CCP-positive, number (%)	63 (80.8)	72 (94.7)	0.013
RF-positive, number (%)	63 (80.8)	65 (85.5)	0.52
CDAI, mean (SD)	21.2 (8.9)	25.0 (11.1)	0.019
High CDAI (>22), number (%)	34 (43.6)	42 (55.3)	0.20
Concurrent MTX use, number (%)	45 (57.7)	39 (51.3)	0.52
Concurrent PSL use, number (%)	19 (24.4)	23 (30.3)	0.47
Therapeutic outcomes at 12 months			
CDAI, mean (SD)‡	11.4 (10.7)	13.9 (11.9)	0.17
Dropout, number (%)	21 (26.9)	23 (30.3)	0.51
Lack or loss of efficacy	14 (17.9)	13 (17.1)	1.00
Adverse events	6 (7.7)	9 (11.8)	0.43
Lost to follow-up	1 (1.3)	1 (1.3)	1.00
Remission (CDAI≤2.8), number (%)	14 (17.9)	10 (13.2)	0.51
Low CDAI (>2.8 and ≤10), number (%)	35 (44.9)	25 (32.9)	0.14
Remission or low CDAI, number (%)	49 (62.8)	35 (46.1)	0.052
High or moderate, number (%)	8 (10.3)	18 (23.7)	0.032
Improvements at 12 months, number (%)§			
CDAI85 (major response)	16 (20.5)	13 (17.1)	0.68
CDAI70 (moderate response)	32 (41.0)	24 (31.6)	0.24
CDAI50 (minor response)	42 (53.8)	40 (52.6)	1.00
MCID-based CDAI improvement	47 (60.3)	41 (53.9)	0.52

*Comparisons of baseline characteristics or CDAI-based improvement measures between patients with one-class failure and those with one-class or two-class failure, using the independent-measures t-test for continuous variables and the χ^2 test for categorical variables.
†Advanced stage was defined as Steinbrocker radiological stages III and IV.
‡To calculate mean CDAI values at 12 months, missing data on dropout patients were imputed using baseline observation carried forward.
§For classification of patients at 12 months, non-responder imputation was used for missing data on patients who had withdrawn from the study because of lack or loss of efficacy, adverse events and lost to follow-up.
anti-CCP, anti-cyclic citrullinated peptide antibodies; bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; MTX, methotrexate; PSL, prednisolone; RA, rheumatoid arthritis; RF, rheumatoid factor.;

and remission during the first 12 months, compared with tocilizumab, for bDMARD-naïve patients. These effects were not observed in previous bDMARD-failure patients. Among tofacitinib-treated patients with previous bDMARD failure, the number of failed bDMARD classes, after adjusted baseline CDAI values, had no negative impacts on any of the CDAI-based improvements. There was no significant difference in the rate of overall adverse events between tofacitinib and tocilizumab treatment for bDMARD-naïve or failure patients.

In the ORAL Standard study, tofacitinib combination therapy with MTX had numerically similar efficacy to adalimumab and MTX combination therapy in patients with active RA with an incomplete response to MTX therapy.[5] In the ORAL Strategy study, tofacitinib and MTX combination therapy was non-inferior to adalimumab and MTX combination therapy in the treatment of patients with RA with an adequate response to MTX.[10] In these studies, most patients were naïve to bDMARDs. These trials suggested that the addition of tofacitinib or adalimumab to MTX therapy is equally efficacious in MTX-inadequate responders. In the present study, approximately one-third of patients did not receive MTX concomitantly, but even after adjusting for the concurrent use of MTX, tofacitinib was more likely to induce and maintain remission, CDAI85 and CDAI70 during the

Table 6 Tofacitinib versus tocilizumab: comparison of CDAI-based improvements at 12 months in patients with RA after IPTW adjustment

	bDMARD-naïve patients with RA		Previous bDMARD-failure patients with RA	
	Adjusted OR* (95% CI)	P value	Adjusted OR* (95% CI)	P value
Remission (CDAI≤2.8)	3.98 (2.22 to 7.15)	<0.001	1.19 (0.55 to 2.55)	0.66
Remission or low CDAI (≤10)	1.92 (1.06 to 3.48)	0.032	1.85 (1.06 to 3.23)	0.031
CDAI85[†] (major response)	4.13 (2.33 to 7.32)	<0.001	0.94 (0.46 to 1.90)	0.86
CDAI70[†] (moderate response)	3.04 (1.73 to 5.35)	<0.001	1.32 (0.73 to 2.38)	0.36
CDAI50[†] (minor response)	1.51 (0.84 to 2.71)	0.17	1.56 (0.89 to 2.72)	0.12
MCID-based improvement[‡]	1.52 (0.84 to 2.75)	0.17	1.68 (0.96 to 2.93)	0.067

*IPTW-adjusted ORs (95% CI) of tofacitinib versus tocilizumab were determined for each of the CDAI-based improvement measures according to univariable logistic regression analyses.
†Defined as achieving and maintaining ≥50% improvement of CDAI (CDAI50), ≥70% (CDAI70) and ≥85% (CDAI85) during the 12-month treatment.
‡Defined as CDAI reduction >12 for patients starting with a high CDAI and CDAI reduction >6 for those starting with a moderate CDAI at 12 months of treatment.
bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; IPTW, inverse probability of treatment weighting; MCID, minimal clinically important difference; RA, rheumatoid arthritis.

first year than tocilizumab, in bDMARD-naïve patients with RA.

In a post hoc study using data from phase II and III trials of tofacitinib, Charles-Schoeman et al[38] showed that the rates of patients achieving ACR criteria-based improvements (ACR20/50/70) at 3 and 6 months were higher in bDMARD-naïve patients versus patients with an inadequate response to bDMARDs. In a real-world cohort study with a mean follow-up of 1.2 years, Mueller et al[39] also showed that the rate of patients who had achieved remission or low disease activity was significantly higher in bDMARD-naïve patients compared with bDMARD-exposed patients. These studies suggested that tofacitinib is more effective in bDMARD-naïve patients compared with previous bDMARD-failure patients. In a post hoc study using data from two randomised clinical trials, Aletaha et al[40] indicated that previous use of csDMARDs, especially the use of MTX plus more than two csDMARDs, affected treatment response to adalimumab at week 24 in patients with established RA regardless of disease duration, and speculated that repeated failure of csDMARDs might select a phenomenon that is resistant to new treatment because different pathogenetic pathways may have become imprinted. In the present study, the number of failed bDMARD classes was not associated with the CDAI-based improvements among tofacitinib-treated previous bDMARD-failure patients. Considering that bDMARDs are more related to RA pathogenesis compared with csDMARDs, even one-class bDMARD failure might have been enough to induce resistance to new treatment.

Table 7 Tofacitinib versus tocilizumab: comparison of CDAI-based improvements at 12 months in patients with RA after IPTW adjustment for time-varying confounders

	bDMARD-naïve patients with RA		Previous bDMARD-failure patients with RA	
	Adjusted OR* (95% CI)	P value	Adjusted OR* (95% CI)	P value
Remission (CDAI ≤2.8)	4.13 (2.18 to 7.82)	<0.001	1.12 (0.51 to 2.41)	0.78
Remission or low CDAI (≤10)	2.57 (1.23 to 5.35)	0.012	1.11 (0.63 to 1.94)	0.73
CDAI85[†] (major response)	4.22 (2.26 to 7.88)	<0.001	0.85 (0.42 to 1.70)	0.64
CDAI70[†] (moderate response)	3.27 (1.72 to 6.23)	<0.001	1.38 (0.76 to 2.50)	0.29
CDAI50[†] (minor response)	1.77 (0.85 to 3.69)	0.13	1.18 (0.67 to 2.08)	0.56
MCID-based improvement[‡]	2.99 (1.19 to 7.50)	0.020	0.79 (0.43 to 1.45)	0.45

*IPTW-adjusted ORs (95% CI) of tofacitinib versus tocilizumab were determined for each of the CDAI-based improvement measures according to univariable logistic regression analyses. A stabilised IPTW was created based on PS estimations at baseline, 3, 6, 9 and 11 months and used as the weight for outcome modelling at 12 months.
†Defined as achieving and maintaining ≥50% improvement of CDAI (CDAI50), ≥70% (CDAI70) and ≥85% (CDAI85) during the 12-month treatment.
‡Defined as CDAI reduction >12 for patients starting with a high CDAI and CDAI reduction >6 for those starting with a moderate CDAI at 12 months of treatment.
bDMARD, biological disease-modifying antirheumatic drug; CDAI, clinical disease activity index; IPTW, inverse probability of treatment weighting; MCID, minimal clinically important difference; RA, rheumatoid arthritis.

In recent phase III trials for other oral JAK inhibitors (the SELECT-COMPARE trial for upadacitinib and the RA-BEAM trial for baricitinib), upadacitinib was superior to adalimumab, in combination with background MTX, for improving signs, symptoms and physical function in patients with RA who had experienced an inadequate response to MTX. In that study, approximately 90% of patients were bDMARD-naïve.[41] Baricitinib in combination with background MTX was also associated with significant clinical improvements as compared with adalimumab plus MTX therapy in bDMARD-naïve patients with active RA who had experienced an inadequate response to MTX.[42]

There are several limitations to this study. First, although subcutaneous injection of tocilizumab every week (QW) has been approved since June 2017 in Japan for treatment of RA with inadequate response to subcutaneous injection every other week,[43 44] no patients participating in this study were treated with the QW dosing regimen of tocilizumab. If we had used the QW in this study, more favourable outcomes might have been induced in the tocilizumab group. Second, this was a multicentre observational study, which might have caused a centre effect. To reduce inter-centre or inter-physician differences, we confirmed that the inclusion/exclusion criteria as well as the administration and follow-up guidelines were abided in all participating institutions. Third, we compared therapeutic outcomes between tofacitinib and tocilizumab treatments in our cohort. If we had included new users of TNF-α inhibitors as the third treatment arm in this study, we could have provided more information on the position of tofacitinib in DMARD therapies of RA. Unfortunately, we did not have a registry that included all patients who had first started TNF-α inhibitors during the same period at the participating institutions. Fourth, we used two imputations (BOCF and NRI) for dropout patients. If the rate of dropout patients due to lost to follow-up had been high, these imputations would have reduced the study's validity. In the present study, however, the rate of lost to follow-up was only 1% of all patients in each treatment group. The rates of dropout patients due to adverse events were less than 7.5% of all patients in each treatment, and half of these patients failed to have controlled disease activity despite tofacitinib or tocilizumab treatment for 3 months or more. The rate of dropout patients due to lost to follow-up, adverse events or lack/loss of efficacy was not significantly different between tofacitinib and tocilizumab treatments. In addition, we confirmed that the results of the sensitivity analysis using MI were consistent with those of the primary analysis based on the NRI and BOCF imputations. Fifth, unlike randomised clinical trials, which theoretically balance known and unknown variables through randomisation, PS-based methods can balance only known co-founders, namely, only the observed confounders that were measured and collected in the study. Finally, the population in this study was small and the results of this study might be different than a true head-to-head study that is properly powered

for both bDMARD-naïve and previous bDMARD-failure patients. We hope that our data will be used to design larger confirmatory studies with increased power.

In conclusion, compared with tocilizumab, tofacitinib induced and maintained CDAI85, CDAI70 and remission at higher rates during the first year in bDMARD-naïve patients with active RA refractory to MTX. These effects were not observed in previous bDMARD-failure patients. There was no significant difference in the rate of discontinuation due to adverse events between both treatments. These findings provide important information that is expected to aid in determining the position of tofacitinib in the treatment algorithm for RA. As current standard-of-care therapy, bDMARDs are preferentially used in patients with RA who have shown an inadequate response to MTX. Given the greater effectiveness of tofacitinib versus tocilizumab in bDMARD-naïve patients, however, tofacitinib could be considered as another option for such patients with RA prior to the initiation of bDMARDs.

Author affiliations
[1]Rheumatology, National Hospital Organisation Kumamoto Saishun Medical Center, Koshi, Japan
[2]Rheumatology, Tsugaru General Hospital United Municipalities of Tsugaru, Goshogawara, Japan
[3]Rheumatology, Yoshitama Clinic for Rheumatic Diseases, Kirishima, Japan
[4]Rheumatic and Collagen Disease Center, Sasebo Chuo Hospital, Sasebo, Japan

ORCID iDs
Shunsuke Mori http://orcid.org/0000-0001-7972-4252
Yukitomo Urata http://orcid.org/0000-0002-5424-2203

REFERENCES

1 Fleischmann R. Tofacitinib in the treatment of active rheumatoid arthritis in adults. *Immunotherapy* 2018;10:39–56.
2 Yamaoka K. Tofacitinib for the treatment of rheumatoid arthritis: an update. *Expert Rev Clin Immunol* 2019;15:577–88.
3 Nash P, Kerschbaumer A, Dörner T, *et al*. Points to consider for the treatment of immune-mediated inflammatory diseases with Janus kinase inhibitors: a consensus statement. *Ann Rheum Dis* 2021;80:71–87.
4 Fleischmann R, Kremer J, Cush J, *et al*. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. *N Engl J Med* 2012;367:495–507.
5 van Vollenhoven RF, Fleischmann R, Cohen S, *et al*. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. *N Engl J Med* 2012;367:508–19.
6 van der Heijde D, Tanaka Y, Fleischmann R, *et al*. Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. *Arthritis Rheum* 2013;65:559–70.
7 Kremer J, Li Z-G, Hall S, *et al*. Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2013;159:253–61.
8 Burmester GR, Blanco R, Charles-Schoeman C, *et al*. Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. *Lancet* 2013;381:451–60.
9 Lee EB, Fleischmann R, Hall S, *et al*. Tofacitinib versus methotrexate in rheumatoid arthritis. *N Engl J Med* 2014;370:2377–86.
10 Fleischmann R, Mysler E, Hall S, *et al*. Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (oral strategy): a phase 3b/4, double-blind, head-to-head, randomised controlled trial. *Lancet* 2017;390:457–68.
11 van der Heijde D, Strand V, Tanaka Y, *et al*. Tofacitinib in combination with methotrexate in patients with rheumatoid arthritis: clinical efficacy, radiographic, and safety outcomes from a twenty-four-month, phase III study. *Arthritis Rheumatol* 2019;71:878–91.
12 Yamanaka H, Tanaka Y, Takeuchi T, *et al*. Tofacitinib, an oral Janus kinase inhibitor, as monotherapy or with background methotrexate, in Japanese patients with rheumatoid arthritis: an open-label, long-term extension study. *Arthritis Res Ther* 2016;18:34.
13 Cohen SB, Tanaka Y, Mariette X, *et al*. Long-term safety of tofacitinib for the treatment of rheumatoid arthritis up to 8.5 years: integrated analysis of data from the global clinical trials. *Ann Rheum Dis* 2017;76:1253–62.
14 Fleischmann R, Wollenhaupt J, Takiya L, *et al*. Safety and maintenance of response for tofacitinib monotherapy and combination therapy in rheumatoid arthritis: an analysis of pooled data from open-label long-term extension studies. *RMD Open* 2017;3:e000491.
15 Wollenhaupt J, Lee E-B, Curtis JR, *et al*. Safety and efficacy of tofacitinib for up to 9.5 years in the treatment of rheumatoid arthritis: final results of a global, open-label, long-term extension study. *Arthritis Res Ther* 2019;21:89.
16 Pope JE, Keystone E, Jamal S, *et al*. Persistence of tofacitinib in the treatment of rheumatoid arthritis in open-label, long-term extension studies up to 9.5 years. *ACR Open Rheumatol* 2019;1:73–82.
17 Mori S, Yoshitama T, Ueki Y. Tofacitinib therapy for rheumatoid arthritis: a direct comparison study between biologic-naïve and experienced patients. *Intern Med* 2018;57:663–70.
18 Mori S, Yoshitama T, Hidaka T, *et al*. Effectiveness and safety of tocilizumab therapy for patients with rheumatoid arthritis and renal insufficiency: a real-life registry study in Japan (the ACTRA-RI study). *Ann Rheum Dis* 2015;74:627–30.
19 Mori S, Yoshitama T, Abe Y, *et al*. Retention of tocilizumab with and without methotrexate during maintenance therapy for rheumatoid arthritis: the ACTRA-RI cohort study. *Rheumatology* 2019;58:1274–84.
20 Arnett FC, Edworthy SM, Bloch DA, *et al*. The American rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
21 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
22 Smolen JS, Landewé RBM, Bijlsma JWJ, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020;79:685–699.
23 Smolen JS, Aletaha D. Interleukin-6 receptor inhibition with tocilizumab and attainment of disease remission in rheumatoid arthritis: the role of acute-phase reactants. *Arthritis Rheum* 2011;63:43–52.
24 Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23:S100–8.
25 Aletaha D, Martinez-Avila J, Kvien TK, *et al*. Definition of treatment response in rheumatoid arthritis based on the simplified and the clinical disease activity index. *Ann Rheum Dis* 2012;71:1190–6.
26 Curtis JR, Yang S, Chen L, *et al*. Determining the minimally important difference in the clinical disease activity index for improvement and worsening in early rheumatoid arthritis patients. *Arthritis Care Res* 2015;67:1345–53.
27 Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat* 2011;10:150–61.
28 Deb S, Austin PC, Tu JV, *et al*. A review of propensity-score methods and their use in cardiovascular research. *Can J Cardiol* 2016;32:259–65.
29 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
30 Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015;34:3661–79.
31 Allan V, Ramagopalan SV, Mardekian J, *et al*. Propensity score matching and inverse probability of treatment weighting to address confounding by indication in comparative effectiveness research of oral anticoagulants. *J Comp Eff Res* 2020;9:603–14.
32 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.
33 Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11:561–70.
34 Mansournia MA, Etminan M, Danaei G, *et al*. Handling time varying confounding in observational research. *BMJ* 2017;359:j4587.
35 Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? *Nephrol Dial Transplant* 2017;32:ii84–90.
36 Fewell Z, Hernán MA, Wolfe F, *et al*. Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004;4:402–20.
37 Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transplant* 2013;48:452–8.
38 Charles-Schoeman C, Burmester G, Nash P, *et al*. Efficacy and safety of tofacitinib following inadequate response to conventional synthetic or biological disease-modifying antirheumatic drugs. *Ann Rheum Dis* 2016;75:1293–301.
39 Mueller RB, Hasler C, Popp F, *et al*. Effectiveness, tolerability, and safety of tofacitinib in rheumatoid arthritis: a retrospective analysis of real-world data from the St. Gallen and Aarau cohorts. *J Clin Med* 2019;8:1548. doi:10.3390/jcm8101548
40 Aletaha D, Maa J-F, Chen S, *et al*. Effect of disease duration and prior disease-modifying antirheumatic drug use on treatment outcomes in patients with rheumatoid arthritis. *Ann Rheum Dis* 2019;78:1609–15.
41 Fleischmann R, Pangan AL, Song I-H, *et al*. Upadacitinib versus placebo or adalimumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase III, double-blind, randomized controlled trial. *Arthritis Rheumatol* 2019;71:1788–800.
42 Taylor PC, Keystone EC, van der Heijde D, *et al*. Baricitinib versus placebo or adalimumab in rheumatoid arthritis. *N Engl J Med* 2017;376:652–62.

43 Ogata A, Tanaka Y, Ishii T, *et al*. Long-Term safety and efficacy of Weekly subcutaneous tocilizumab monotherapy in patients with rheumatoid arthritis who had an inadequate response to subcutaneous tocilizumab every other week: results from the open-label extension of the SHINOBI study. *Mod Rheumatol* 2019;29:767–74.

44 Ogata A, Tanaka Y, Ishii T, *et al*. A randomized, double-blind, parallel-group, phase III study of shortening the dosing interval of subcutaneous tocilizumab monotherapy in patients with rheumatoid arthritis and an inadequate response to subcutaneous tocilizumab every other week: results of the 12-week double-blind period. *Mod Rheumatol* 2018;28:76–84.