



OPEN ACCESS

CLINICAL SCIENCE

Efficacy and safety of tofacitinib versus baricitinib in patients with rheumatoid arthritis in real clinical practice: analyses with propensity score-based inverse probability of treatment weighting

Yusuke Miyazaki ,¹ Kazuhisa Nakano,¹ Shingo Nakayamada,¹ Satoshi Kubo ,¹ Yoshino Inoue,¹ Yoshihisa Fujino,² Yoshiya Tanaka ¹**Handling editor** Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/annrheumdis-2020-219699>).

¹The First Department of Internal Medicine, University of Occupational and Environmental Health, Japan, Kitakyushu, Japan

²Department of Environmental Epidemiology, University of Occupational and Environmental Health, Japan, Kitakyushu, Japan

Correspondence to

Professor Yoshiya Tanaka, First Department of Internal Medicine, University of Occupational and Environmental Health Japan, Kitakyushu 807-8555, Japan; tanaka@med.uoeh-u.ac.jp

Received 12 December 2020

Revised 22 March 2021

Accepted 25 March 2021

Published Online First

7 April 2021

ABSTRACT

Objectives The differences of efficacy between each Janus kinase (JAK) inhibitors have not been clarified in the patients with rheumatoid arthritis (RA) in clinical practice. Here, we compared the efficacy between tofacitinib (TOFA) and baricitinib (BARI) in clinical practice.

Methods The efficacy of TOFA (n=156) in patients with RA was compared with BARI (n=138). Selection bias was reduced to a minimum using propensity score-based inverse probability of treatment weighting (IPTW). The Clinical Disease Activity Index (CDAI) trajectory for patients who started TOFA or BARI was analysed using growth mixture modelling (GMM).

Results No significant difference was observed in patient characteristics between the TOFA and BARI groups in after adjustment by propensity score-based IPTW. The BARI group had a significantly higher rate of CDAI remission at week 24 after the introduction of JAK inhibitors than the TOFA group. The treatment-resistant group defined by GMM, comprising patients who did not achieve low disease activity at week 24, was more likely to include those who had received many biological disease-modifying antirheumatic drugs (bDMARDs) before the introduction of JAK inhibitors and those who received TOFA. Among patients with RA who received TOFA, those who had received ≥ 4 bDMARDs before the introduction of TOFA were more likely to be classified into the treatment-resistant group.

Conclusions BARI showed a similar safety profile and better clinical outcome when compared with TOFA after reduction to a minimum of selection bias. However, these were observed in a small population. Accordingly, further investigation is required in an accurately powered head-to-head trial.

INTRODUCTION

Rheumatoid arthritis (RA) is a systemic inflammatory disease that causes progressive bone and joint destruction and irreversible physical dysfunction.^{1–3} In the last 20 years, a paradigm shift has occurred in the treatment of RA with the advent of biological disease-modifying antirheumatic drugs (bDMARDs).⁴ However, owing to their high molecular weight, bDMARDs can be administered only via the parenteral route and are associated with secondary failure.⁵ To address these issues, Janus kinase (JAK) inhibitors, which are orally

Key messages**What is already known about this subject?**

- Randomised controlled trials have confirmed the efficacy of tofacitinib (TOFA) and baricitinib (BARI) monotherapies in patients with rheumatoid arthritis (RA) who are methotrexate (MTX) naïve and those who achieve MTX-inadequate response (IR), tumour necrosis factor inhibitor-IR and biological disease-modifying antirheumatic drug (bDMARD) IR.

What does this study add?

- When the Clinical Disease Activity Index (CDAI) after 24 weeks of treatment was compared between TOFA or BARI after reduction of the selection bias to a minimum and adjustment for patient characteristics by propensity score-based inverse probability of treatment weighting, BARI was more effective.
- Trajectory analysis of the changes in CDAI for TOFA and BARI divided the patients into three trajectory groups.
- Among the three groups was a treatment-resistant group that did not achieve low disease activity at week 24 after the introduction of Janus kinase inhibitors and was more likely to include patients with RA treated with TOFA, particularly those resistant to multiple bDMARDs.

How might this impact on clinical practice or future developments?

- Results suggest that TOFA may be less effective in patients resistant to multiple bDMARDs, while BARI may be more effective after 24 weeks of treatment.

administered low molecular weight compounds, have been used. Among the JAK inhibitors currently available, tofacitinib (TOFA) and baricitinib (BARI) have been widely used in many regions for RA treatment. TOFA is a selective inhibitor of JAK1 and JAK3, and its inhibitory effect on JAK2 and tyrosine kinase (TYK) 2 is limited.⁶ BARI is a selective inhibitor of JAK1 and JAK2 and exhibits a moderate inhibitory activity against TYK2, while its inhibitory activity against JAK3 is limited.⁷ Randomised



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Miyazaki Y, Nakano K, Nakayamada S, et al. *Ann Rheum Dis* 2021;**80**:1130–1136.

controlled trials have shown that TOFA and BARI monotherapies are effective in patients with RA who are methotrexate (MTX) naïve^{8 9} or have MTX-inadequate response (IR),^{10 11} tumour necrosis factor inhibitor-IR^{12 13} and bDMARDs-IR.¹⁴ In vitro studies have revealed variations in the pharmacological effects of JAK inhibitors at the cellular level.^{15 16} However, such variations have not been investigated in real-world clinical practice. Therefore, the selection of JAK inhibitor for RA treatment based on patient type remains a major concern.

In the present study, we compared the efficacy and safety of TOFA and BARI in real-world clinical practice after reduction to a minimum of the selection bias, using propensity score-based inverse probability of treatment weighting (IPTW) and adjustment for confounding patient characteristics. Growth mixture modelling (GMM) is a method of analysis to identify trajectory groups into which longitudinal changes in factors can be classified.¹⁷ This method allows the identification of characteristics of each trajectory group and analysis of factors affecting trajectories. We analysed the trajectories of changes in disease activity in patients receiving TOFA or BARI using GMM and evaluated clinical characteristics of their responses to both the drugs.

MATERIALS AND METHODS

Patients and study design

Patients were recruited from the FIRST registry, a registry study of patients with RA receiving molecularly targeting antirheumatic drugs at multiple institutions affiliated to our university hospital, the key station.^{18–20} RA was diagnosed when patients met the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism classification criteria or the 1987 ACR classification criteria.^{21 22} The observation period of the study was 24 weeks.

Treatment with JAK inhibitors

TOFA or BARI was administered to patients with RA in whom disease activity could not be controlled by standard doses of MTX or conventional synthetic DMARDs (csDMARDs) or in patients with RA for whom csDMARDs, including MTX, could not be used. There was no major difference observed in the proportion of patients who were allocated to TOFA or BARI at each site; no major difference was noted in the selection of JAK inhibitors at each site. Dose of JAK inhibitor was shown in online supplemental material.

Clinical efficacy and outcome

The primary outcome was rate of remission at week 24 in each group, measured by the Clinical Disease Activity Index (CDAI).^{23 24} CDAI remission was defined as a score of ≤ 2.8 and low disease activity (LDA) was defined as a score of ≤ 10.0 . Additional secondary outcomes included disease activity, retention rate and safety at week 24. The analyses were performed by the last observation carried forward (LOCF) and non-responder imputation (NRI) was also used to evaluate CDAI, Simplified Disease Activity Index (SDAI) and Health Assessment Questionnaire (HAQ) remission rates.

Safety

The incidence and severity of all adverse events were recorded. The National Institutes of Health Common Terminology Criteria for Adverse Events (V.3.0) were used to describe adverse events and laboratory abnormalities.

Propensity score-based IPTW

To adjust for baseline patient characteristics between the two groups, the calculated propensity scores were weighted using the ‘ratio of patients receiving BARI to all patients/propensity score’ in the BARI group and the ‘ratio of patients receiving TOFA to all patients/1 propensity score’ in patients treated with TOFA as the weighting coefficient on stability. Details of the procedure of calculating propensity score are shown in online supplemental material.

Growth mixture modelling

To understand patient response patterns after receiving TOFA and BARI, GMM was applied to classify patients into different subgroups based on CDAI trajectories.²⁵ GMM was performed with STATA V.16.0 (StataCorp LLC, College Station, Texas, USA).¹⁷ Details of the procedure of GMM are shown in online supplemental material.

Other statistical analyses

Patient characteristics are expressed as mean \pm SD, median (IQR) or number (%) of patients. Kaplan-Meier method was used to assess the retention rates, and the differences between the TOFA and BARI groups were analysed by the log-rank test. Student’s t-test, Mann-Whitney’s U test or one-way analysis of variance (ANOVA) were used for between-group comparisons, and the Pearson’s χ^2 test was used for the comparison of categorical variables. All reported p values are two-sided and were not adjusted for multiple testing. The level of significance was $p < 0.05$. All analyses were conducted using JMP V.14.0 (SAS Institute, Cary, North Carolina, USA) and SPSS software V.25.0.

Table 1 Safety and laboratory data, weeks 0–24

| Variables | TOFA, n=156 | BARI, n=138 | P value |
|--|------------------|------------------|---------|
| Safety data | | | |
| Serious adverse events, n (%) | 10 (6.4%) | 4 (2.9%) | 0.15 |
| Any adverse event after start of therapy, n (%) | 56 (35.9%) | 31 (22.5%) | 0.04 |
| Infection, n (%) | 37 (23.7%) | 23 (16.7%) | 0.13 |
| Herpes zoster, n (%) | 2 (1.3%) | 5 (3.6%) | 0.18 |
| Serious infection, n (%) | 2 (1.3%) | 2 (1.5%) | 0.90 |
| Cancer, n (%) | 0 (0.0%) | 0 (0.0%) | 1.00 |
| Major adverse cardiovascular event, n (%) | 0 (0.0%) | 0 (0.0%) | 1.00 |
| Venous thromboembolism, n (%) | 0 (0.0%) | 0 (0.0%) | 1.00 |
| Laboratory data—median change from baseline | | | |
| Haemoglobin (g/L) | 1.5 (–4.0–8.8) | 0.5 (–7.0–8.0) | 0.21 |
| Neutrophils (μ L) | –857 (–2057–268) | –817 (–2170–47) | 0.67 |
| Lymphocytes (μ L) | 111 (–172–458) | 232 (–134–621) | 0.10 |
| Alanine aminotransferase (IU/L) | 2 (–4–9) | 5 (–1–10) | 0.09 |
| Creatinine (mg/dL) | 0.06 (0.00–0.12) | 0.08 (0.03–0.13) | 0.11 |
| Creatine phosphokinase (IU/L) | 30 (4–65) | 39 (4–70) | 0.51 |

Adverse events, infection or laboratory abnormalities leading to permanent discontinuation of the JAK inhibitor are designated as serious adverse events. The data shown are numbers and percentages of patients with adverse events. Laboratory values are reported as the least-squares-mean change from baseline at week 24. BARI, baricitinib; CDAI, Clinical Disease Activity Index; JAK, Janus kinase; TOFA, tofacitinib.

Table 2 Patient characteristics in the TOFA and BARI groups before and after IPTW

| Variables | Before IPTW | | | After IPTW | | |
|---|------------------------------------|-----------------------------------|---------|-----------------------------------|-----------------------------------|---------|
| | TOFA, n=156 | BARI, n=138 | P value | TOFA, n=153* | BARI, n=141* | P value |
| Age (years) | 58.9±13.2 | 57.2±13.6 | 0.25 | 58.2±13.4 | 58.2±13.3 | 0.96 |
| Sex, n (% female) | 129 (82.7) | 109 (79.0) | 0.16 | 126 (82.4) | 115 (81.6) | 0.86 |
| Disease duration (month) | 96 (35–192) | 77 (24–158) | 0.16 | 118.5±103.7 | 122.3±120.0 | 0.77 |
| Treatment history | | | | | | |
| MTX use at baseline, n (%) | 117 (75.0) | 94 (68.1) | 0.20 | 112 (73.2) | 103 (73.1) | 0.53 |
| Dose, mg/w | 12.3±3.5 | 11.7±3.7 | 0.18 | 12.2±3.5 | 12.0±3.6 | 0.68 |
| Glucocorticoid use at baseline, n (%) | 20 (12.8) | 33 (23.9) | 0.02 | 25 (16.3) | 24 (17.0) | 0.88 |
| Dose, mg/day | 5.0 (2.5–7.5) | 7.5 (5.0–10.0) | 0.14 | 5 (2.5–6.4) | 7.5 (2.5–10.0) | 0.26 |
| bDMARD naïve, n (%) | 37 (23.7) | 45 (32.6) | 0.09 | 43 (28.1) | 40 (28.4) | 0.52 |
| Number of previous bDMARDs use, n 1/2/3/4/≥5 | 38/30/26/19/6 | 37/22/17/8/9 | 0.24 | 36/27/25/17/5 | 38/24/20/9/10 | 0.60 |
| JAK inhibitor dose, n (%) | 10 mg=140 (89.7) 5 mg=16 (10.3) | 4 mg=122 (88.4) 2 mg=16 (11.6) | | 10 mg=135 (90.8) 5 mg=10 (9.2) | 4 mg=122 (86.5) 2 mg=19 (13.5) | |
| 28-tender joint count | 9.2±6.1 | 9.2±6.7 | 0.98 | 9.2±6.2 | 9.0±6.5 | 0.72 |
| 28-swollen joint count | 7.0±4.6 | 7.8±5.8 | 0.21 | 7.3±4.6 | 7.3±5.3 | 0.89 |
| GH, VAS 0–100 mm | 53.3±24.1 | 52.8±24.6 | 0.86 | 52.9±23.3 | 52.7±24.3 | 0.95 |
| EGA, VAS 0–100 mm | 44.9±20.8 | 47.4±22.0 | 0.31 | 45.9±20.7 | 45.6±21.4 | 0.92 |
| Pain VAS 0–100 mm | 52.4±24.6 | 51.1±26.6 | 0.67 | 52.2±24.0 | 51.8±25.6 | 0.91 |
| DAS28-ESR | 5.3±1.3 | 5.2±1.3 | 0.34 | 5.4±1.3 | 5.2±1.3 | 0.29 |
| SDAI | 27.6±12.9 | 27.5±13.3 | 0.94 | 27.8±12.8 | 26.9±12.6 | 0.53 |
| CDAI | 25.8±11.7 | 26.1±12.7 | 0.88 | 26.3±11.7 | 25.4±11.9 | 0.81 |
| HAQ-DI | 1.3±0.8 | 1.2±0.8 | 0.12 | 1.3±0.8 | 1.2±0.7 | 0.90 |
| EQ-5D | 0.6±0.1 | 0.6±0.1 | 0.43 | 0.6±0.1 | 0.6±0.1 | 0.86 |
| CRP (mg/dL) | 0.4 (0.1–1.8) | 0.4 (0.1–1.3) | 0.64 | 0.4 (0.1–1.7) | 0.4 (0.1–1.6) | 0.76 |
| ESR (mm/hour) | 39.4±30.1 | 38.2±30.1 | 0.73 | 40.4±31.5 | 39.4±30.9 | 0.80 |
| Rheumatoid factor positive, n (%) | 121 (77.6) | 107 (77.5) | 1.00 | 118 (77.1) | 107 (75.9) | 0.89 |
| Rheumatoid factor (U/mL) | 69.2 (18.6–157.9) | 51.4 (13.7–150.7) | 0.39 | 70.3 (18.1–170.4) | 46.5 (11.4–121.9) | 0.85 |
| Anti-CCP antibody, n (%) | 118 (75.6) | 98 (71.0) | 0.43 | 115 (75.2) | 99 (70.2) | 0.36 |
| Anti-CCP antibody (U/mL) | 41.7 (5.2–265.8) | 76.2 (2.6–397.1) | 0.56 | 45.2 (5.3–272.1) | 70.6 (1.9–386.4) | 0.78 |
| MMP-3 (ng/mL) | 93 (55–264) | 118 (51–229) | 0.82 | 86 (53–252) | 111 (50–234) | 0.80 |

Data are mean±SD, median (IQR) or number (%) of patients.

*The number of subjects changed after IPTW in the calculation; however, the actual number of subjects did not change.

BARI, baricitinib; bDMARDs, biological disease-modifying antirheumatic drugs; CCP, cyclic citrullinated peptide; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; DAS, Disease Activity Score; EGA VAS, Evaluator Global Assessment of Disease Activity Visual Analogue Scale; EQ-5D, EuroQol 5 dimension; ESR, erythrocyte sedimentation rate; HAQ-DI, Health Assessment Questionnaire-Disability Index; IPTW, inverse probability of treatment weighting; JAK, Janus kinase; MMP-3, matrix metalloproteinase 3; MTX, methotrexate; SDAI, Simplified Disease Activity Index; TOFA, tofacitinib; GH VAS, patient’s global assessment of disease activity visual analogue scale.

RESULTS

Comparison of the efficacy and safety of TOFA and BARI

Totally, 156 patients with RA who were followed-up for ≥6 months after the introduction of TOFA and 138 patients who were followed up for ≥6 months after the introduction of BARI between August 2014 and May 2020 were included. Online supplemental figure S1A shows the retention rates over 24 weeks after the introduction of JAK inhibitors in both the groups. There was no significant difference in the retention rate between the two groups (TOFA vs BARI=87.8% vs 91.3%, p=0.31). Table 1 shows the observed adverse events. Although the incidence of ≤grade 2 adverse events, as specified by the National Institutes of Health Common Terminology Criteria for Adverse Events (V3.0), was significantly lower in the BARI group, there was no difference between the two groups in the incidence of serious infections or adverse events that could lead to discontinuation of JAK inhibitors. Laboratory data showed significant decrease in neutrophil count and significant increase in lymphocyte count, creatinine level and creatinine phosphokinase level in both the groups. A

comparison of the changes in laboratory data between the two groups showed no significant differences.

Comparison of the efficacy of TOFA and BARI for 24 weeks after treatment introduction showed that the BARI group had a significantly lower CDAI (TOFA vs BARI=8.2 vs 6.2, p=0.04) and a significantly higher rate of CDAI-LDA (TOFA vs BARI=69.2% vs 81.2%, p=0.02) at week 24 after treatment introduction (online supplemental figure S1B, S1C). No difference was observed in CDAI remission.

Patient characteristics in the TOFA and BARI groups after adjustment by propensity score-based IPTW

Table 2 (left-hand side) shows the patient characteristics before adjustment. The rate of concomitant glucocorticoid (GC) use was significantly lower in the TOFA group than in the BARI group. The TOFA group also included more bDMARDs-naïve patients than the BARI group. Next, we calculated the IPTW using the propensity scores to reduce the selection bias to a minimum and adjusted the patient characteristics. The adjusted patient characteristics are shown in table 2 (right-hand side). No significant

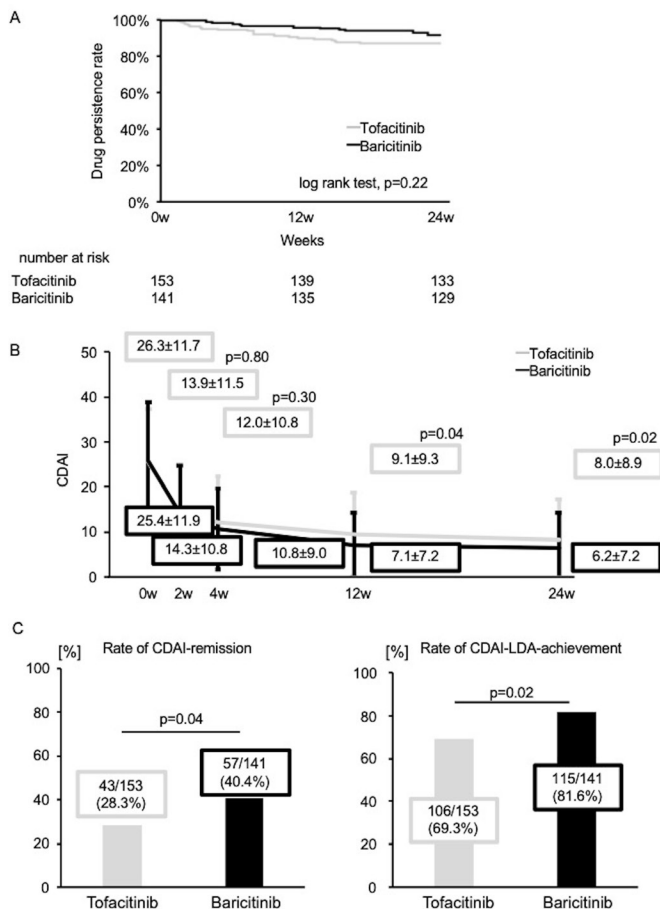


Figure 1 Changes in disease activity over 24 weeks after the introduction of JAK inhibitors after adjustment by propensity score-based IPTW. The selection bias was adjusted by propensity score-based IPTW in patients with rheumatoid arthritis treated with TOFA or BARI. (A) Retention rates over 24 weeks after the introduction of JAK inhibitors (Kaplan-Meier curves). (B) Changes in CDAI over 24 weeks after the introduction of JAK inhibitors: comparison between the TOFA and BARI groups with mean±SD and p values derived from Student's t-test. (C) Comparison of rates of CDAI remission (left) and CDAI-LDA achievement (right) between the two groups by Pearson's χ^2 test. Numbers represent percentages of all patients (%). BARI, baricitinib; CDAI, Clinical Disease Activity Index; IPTW, inverse probability of treatment weighting; JAK, Janus kinase; LDA, low disease activity; TOFA, tofacitinib.

differences were observed in any patient characteristic, and the standardised differences were <0.1 for all the characteristics. The distribution of variables was well balanced.

Comparison of efficacy and safety between the TOFA and BARI groups after adjustment by propensity score-based IPTW

Figure 1 shows the retention rate and efficacy over 24 weeks of treatment with TOFA and BARI after adjustment by IPTW. The retention rates over 24 weeks did not differ between the TOFA and BARI groups (TOFA vs BARI=86.9% vs 91.5%, $p=0.22$) (figure 1A). Adverse events that led to discontinuation of JAK inhibitors are shown in online supplemental table S1. No difference was observed in the incidence of adverse events leading to discontinuation of JAK inhibitors in the TOFA and BARI groups. CDAI at 24 weeks after the introduction of JAK inhibitors was 8.0 ± 8.9 and 6.2 ± 7.2 in the TOFA and BARI

groups, respectively (figure 1B). CDAI, SDAI, HAQ-DI and C reactive protein (CRP) level were significantly improved in both the groups at week 2 after the introduction of JAK inhibitors and further improved until week 24 (table 3).

Compared with the TOFA group using a generalised linear model, the BARI group showed a significantly lower CDAI (Δ CDAI = -1.9 , 95% CI: -3.7 to -0.3 , $p=0.02$) and a significantly higher rate of CDAI remission (OR: 1.7, 95% CI: 1.1 to 2.7, $p=0.04$) at 24 weeks (figure 1C). Similarly, at week 24, SDAI was significantly lower in the BARI group, and the rates of SDAI remission and SDAI-LDA achievement were significantly higher in the BARI group (online supplemental figure S2). Furthermore, no differences were observed in HAQ-DI or rate of HAQ-DI-remission at week 24 (online supplemental figure S3).

Trajectories of changes in CDAI in the TOFA and BARI groups using GMM

Next, we analysed the trajectories of changes in CDAI in 294 patients receiving TOFA or BARI and the differences in changes in CDAI between the TOFA and BARI groups by using GMM. The cubic function-based linear model of trajectory showed the best fit (online supplemental table S2). As for the number of trajectory groups, the best fit was obtained when the patients were divided into the following three groups (online supplemental table S3): group 1 comprising patients with moderate disease activity (MDA) at baseline who exhibited improvement in disease activity immediately after the introduction of JAK inhibitors and achieved LDA at week 24, group 2 comprising patients with high disease activity (HDA) at baseline who exhibited improvement in disease activity immediately after the introduction of JAK inhibitors and achieved LDA at week 24 and group 3 (treatment-resistant group) comprising patients with HDA at baseline who exhibited a partial or limited response to JAK inhibitors after introduction and did not achieve LDA at week 24 (figure 2A and online supplemental table S4).

When the proportion of patients in each trajectory group was compared between the TOFA and BARI groups, the proportion of patients classified as the treatment-resistant group was lower in the BARI group than in the TOFA group (TOFA:BARI=23.7%:13.0%, $p=0.02$) (figure 2B). No difference was observed in the proportion of patients classified as group 1 (TOFA:BARI=50.6%:52.9%, $p=0.70$) and group 2 (TOFA:BARI=25.6%:34.1%, $p=0.11$).

The CDAI improvement rate was analysed, using GMM, and the subjects were divided into groups that followed four trajectories (online supplemental figure S4 and online supplemental table S6-S8). Group B (CDAI improvement rates increased at 12 weeks and maintained an increasing trend until 24 weeks) included a large percentage of subjects belonging to the BARI group. Moreover, Group C (CDAI had improved to approximately half at 24 weeks) included a large percentage of subjects belonging to the TOFA group. Patients who belonged to Group D, among the patients who had used TOFA, included a few bio-naïve patients and many patients who failed to respond to many bDMARDs.

Factors associated with treatment resistance in the TOFA and BARI groups

Multivariable logistic regression analysis was performed to identify factors contributing to belonging to treatment-resistance group (online supplemental table S5, table 4). The explanatory variables were age, female sex, duration of RA, concomitant

Table 3 Change in efficacy 2 weeks, 12 weeks and 24 weeks after the introduction of JAK inhibitors

| | TOFA (n=153)† | | | BARI (n=141)† | | |
|--------------|----------------------|---------------------|---------------------|----------------------|---------------------|---------------------|
| | Change from baseline | | | Change from baseline | | |
| | Week 2 | Week 12 | Week 24 | Week 2 | Week 12 | Week 24 |
| CDAI | -12.3 (11.2)* | -17.2 (12.6)* | -18.1 (13.0)* | -11.5 (11.6)* | -18.3 (13.0)* | -19.3 (14.1)* |
| SDAI | -12.5 (11.6)* | -17.9 (13.5)* | -19.0 (14.1)* | -12.3 (12.5)* | -19.1 (13.8)* | -20.4 (15.0)* |
| HAQ-DI | -0.22 (0.41)* | -0.41 (0.53)* | -0.45 (0.61)* | -0.15 (0.38)* | -0.31 (0.55)* | -0.39 (0.65)* |
| CRP, mg/dL | -0.06 (-0.92–0.00)* | -0.05 (-1.30–0.02)* | -0.11 (-1.30–0.00)* | -0.07 (-0.70–0.00)* | -0.05 (-1.11–0.00)* | -0.13 (-1.55–0.02)* |
| ESR, mm/hour | -5.46 (14.21)* | -7.66 (21.9)* | -10.78 (26.7)* | -6.48 (14.66)* | -10.23 (22.77)* | -12.84 (33.39)* |

Change from baseline data is mean (SD) and median (IQR).

*P<0.001 from within-group mean change from baseline.

†The number of subjects changed after IPTW in the calculation; however, the actual number of subjects did not change.

BARI, baricitinib; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; HAQ-DI, Health Assessment Questionnaire-Disability Index; JAK, Janus kinase; SDAI, Simplified Disease Activity Index; TOFA, tofacitinib.

MTX dose, number of bDMARDs used before JAK inhibitors, TOFA use, HAQ-DI, CRP, matrix metalloproteinase 3, rheumatoid factor (RF) and anti-cyclic citrullinated peptide (CCP) antibody. EuroQol-5 Dimension was excluded from the explanatory

variables because of the collinearity with HAQ-DI. Additionally, CDAI was excluded from the explanatory variables because grouping was based on the trajectories of CDAI.

For all patients receiving JAK inhibitors, the factors contributing to belonging to treatment-resistance group were: high baseline HAQ-DI score (p=0.02) and high number of bDMARDs used before JAK inhibitors (p=0.002) and TOFA use (p=0.03).

When multivariable logistic regression analysis was separately performed for each treatment group, patients receiving more bDMARDs before the JAK inhibitor were more likely to belong to treatment-resistance group in the TOFA group (p<0.001). In the BARI group, multivariable logistic regression analysis did not identify any factors associated with belonging to treatment-resistance group.

DISCUSSION

In the present study, we compared the efficacy and safety at week 24 after the introduction of TOFA and BARI in patients with RA after reducing the selection bias to a minimum using the propensity score-based IPTW. Although the incidence of adverse events was comparable between the two groups, the BARI group showed a significantly lower CDAI and a significantly higher rate of CDAI remission at week 24. Although the CDAI numerical values displayed statistical differences, the differences in numerical values were small and may not be clinically meaningful. Although no differences were observed in HAQ-DI, even up to 24 weeks later, the duration of the analysis might have been too short for differences to be observed.

There are some reports on network meta-analysis indirectly comparing efficacy and safety of TOFA and BARI. Regarding efficacy, some reports suggested that BARI at a dose of 4 mg/day may be more effective than TOFA at a dose of 5 mg/day,^{26 27} whereas another study showed that TOFA at a dose of 10 mg/day may be more effective than BARI at a dose of 4 mg/day.²⁸ In terms of safety, no consistent results have been reported regarding which drug is superior.^{26 28 29} While these reports describe network meta-analysis by indirectly comparing results of randomised controlled trials, to the best of our knowledge, there is no study comparing efficacy and safety in real-world clinical practice. The present study is the first to compare the efficacy of TOFA and BARI in real-world clinical practice.

In this study, patients were divided into two groups: patients with MDA to HDA at baseline (group 1) and patients with HDA at baseline than group 1 (groups 2 and 3) based on the analysis of the trajectories of CDAI using GMM. In groups 1 and 2, disease activity was improved immediately after the introduction of JAK inhibitors. In group 3, disease activity was partially improved,

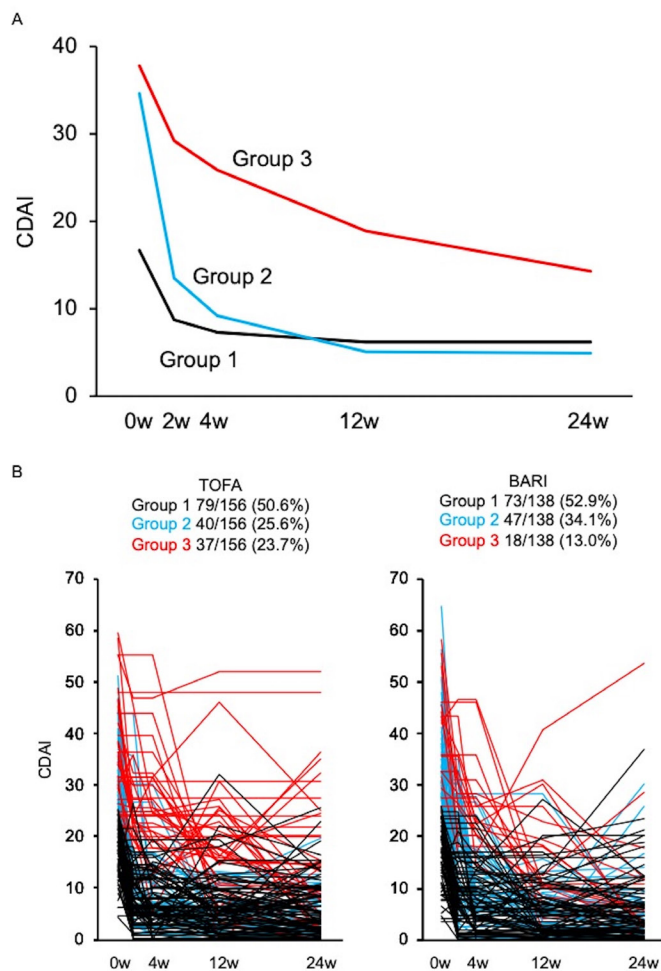


Figure 2 Patient CDAI responses and locally estimated scatterplot smoothing trajectory group modelling for patients receiving TOFA and BARI. (A) Lines are locally estimated scatterplot smoothing trajectories for the three patient trajectory groups. (B) Changes in CDAI in all patients receiving JAK inhibitors (TOFA left) and BARI (right) and the proportions of patients in each trajectory group. Group 1: black line, Group 2: blue line and group 3: red line. BARI, baricitinib; CDAI, Clinical Disease Activity Index; JAK, Janus kinase; TOFA, tofacitinib.

Table 4 Factors for belonging to treatment-resistance group identified by univariable and multivariable logistic regression analyses by treatment group

| | TOFA (n=156) | | | | BARI (n=138) | | | |
|---------------------------------|----------------------|---------|------------------------|---------|----------------------|---------|------------------------|---------|
| | Univariable analysis | | Multivariable analysis | | Univariable analysis | | Multivariable analysis | |
| | OR (95% CI) | P value | OR (95% CI) | P value | OR (95% CI) | P value | OR (95% CI) | P value |
| Age | 1.03 (0.99–1.06) | 0.05 | 1.01 (0.98–1.05) | 0.49 | 1.02 (0.98–1.06) | 0.38 | | |
| Sex (female) | 0.87 (0.33–2.24) | 0.77 | | | 0.92 (0.28–3.04) | 0.89 | | |
| RA duration | 1.00 (0.99–1.01) | 0.50 | 1.00 (0.99–1.01) | 0.11 | 1.00 (0.99–1.00) | 0.85 | 1.00 (0.99–1.01) | 0.24 |
| MTX dose | 0.94 (0.89–0.99) | 0.04 | 0.97 (0.90–1.04) | 0.40 | 1.00 (0.92–1.07) | 0.84 | | |
| Number of previous bDMARDs used | 1.60 (1.20–2.07) | <0.001 | 1.77 (1.26–2.48) | <0.001 | 1.31 (0.97–1.78) | 0.08 | 1.41 (0.95–2.10) | 0.09 |
| HAQ-DI | 2.26 (1.36–3.77) | 0.001 | 1.86 (1.02–3.39) | 0.04 | 2.34 (1.19–4.59) | 0.01 | | |
| CRP | 1.13 (1.01–1.27) | 0.04 | 1.10 (0.97–1.26) | 0.15 | 1.11 (0.97–1.27) | 0.13 | | |
| MMP-3 | 1.00 (0.99–1.00) | 0.14 | | | 1.00 (0.99–1.00) | 0.26 | | |
| Rheumatoid factor titre | 1.00 (0.99–1.00) | 0.63 | | | 1.00 (0.99–1.00) | 0.31 | | |
| Anti-CCP antibody titre | 1.00 (0.99–1.00) | 0.96 | 1.00 (0.99–1.01) | 0.23 | 1.01 (1.00–1.02) | 0.03 | 1.00 (0.99–1.01) | 0.09 |

BARI, baricitinib; bDMARDs, biological disease-modifying antirheumatic drugs; CCP, cyclic citrullinated peptide; CRP, C reactive protein; HAQ-DI, Health Assessment Questionnaire Disability Index; MMP-3, matrix metalloproteinase 3; MTX, methotrexate; RA, rheumatoid arthritis; TOFA, tofacitinib.

and LDA was not achieved at week 24 after the introduction of JAK inhibitors. The patients in group 3 were resistant to treatment. We also performed multivariable logistic regression analysis separately in the TOFA and BARI groups to analyse factors contributing to treatment resistance (group 3). In the TOFA group, patients who had received more bDMARDs before the JAK inhibitor were more likely to be resistant to treatment. We performed logistic regression analysis with the classification of group 3 as the dependent variable and the number of bDMARDs used before JAK inhibitors as the explanatory variable. Even though we similarly analysed the trajectory of CDAI improvement rates using GMM, we found that the subjects using TOFA belonging to the group with the lowest improvement rate included a large percentage of patients who had failed to respond to many bDMARDs. Then, we constructed receiver operating characteristic (ROC) curves to calculate the cut-off value. Results showed that patients receiving ≥ 4 bDMARDs were more likely to be resistant to treatment (sensitivity=0.62, specificity=0.86 and area under the curve=0.77) (data not shown). This suggested that TOFA might be partially effective in patients who received ≥ 4 bDMARDs. In the BARI group, high levels of HAQ-DI and anti-CCP antibody were extracted through univariable analysis as factors likely to belong to the treatment-resistance group, whereas no such factors were extracted following multivariable analysis.

Patients receiving many bDMARDs before JAK inhibitors and those receiving TOFA were more likely to be classified into the treatment resistant group in which CDAI changed, as observed in group 3. Because TOFA was partially effective in patients who did not respond to ≥ 4 bDMARDs, results from the present study might suggest that the efficacy of TOFA differs from that of BARI.

The present study has several limitations. First, this analysis was performed in a small number of Japanese patients, and hence, our findings may not be applicable to all patients with RA. Second, although propensity score-based IPTW was used to reduce the selection bias to a minimum and to adjust patient characteristics, not all confounding factors were adjusted. There may be unknown confounding factors. Third, there is the possibility that bias was introduced because of the use of LOCF. However, missing values were found in only seven patients, and no differences were seen in the results, even after sensitivity analyses

were performed. Moreover, when NRI was used to evaluate the CDAI, SDAI and HAQ remission rates, the BARI group had significantly higher CDAI, SDAI and HAQ remission rates than the TOFA group. Fourth, because the observation period was just 24 weeks, long-term variation in efficacy of TOFA and BARI is not known; particularly, whether the difference revealed in the present study affect bone destruction was unclear. Fifth, because of the small number of patients resistant to treatment with BARI in GMM, we might have been unable to identify factors contributing to treatment resistance by performing multivariable logistic regression analysis. However, univariable analysis also showed that the number of bDMARDs used before BARI was not associated with treatment resistance. In other words, the difference in efficacy due to the number of bDMARDs used before JAK inhibitors might have contributed to the difference in efficacy between TOFA and BARI. As the treatment-resistant group was identified by the analysis using GMM, a model with better fit may be developed by conducting long-term studies. Fifth, in Japan, sales of TOFA began 4 years earlier than BARI, which might have led to a selection bias. However, even if we conducted a similar comparison of efficacy between patients to whom TOFA had been introduced after the date when the use of BARI was allowed in Japan, and patients who had used BARI, we found that the BARI group had higher efficacy (data not shown). Finally, there is no definitive basic study that supports the difference in efficacy between TOFA and BARI. Because the safety and efficacy features identified by basic analysis of signal transduction are not necessarily consistent with those observed in clinical practice, the efficacy of the drugs might have differed in the present study. Thus, further investigation is needed in this regard.

In summary, even if IPTW is used, there is a possibility that the selection bias cannot be removed entirely, and that there are confounding factors that have not been measured. TOFA may be partially effective in patients resistant to many bDMARDs. Consequently, efficacy may differ between TOFA and BARI. TOFA is likely to be less effective in patients with RA resistant to numerous bDMARDs. These results were observed in a relatively small group of patients and were obtained on hypothesis testing; accordingly, they need to be investigated in an accurately powered head-to-head trial.

Acknowledgements The authors thank all medical staff at all participating institutions for providing the data, especially Ms. Hiroko Yoshida, Ms. Youko Saitou, Ms. Machiko Mitsui and Ms. Ayumi Maruyama for the excellent data management in the FIRST registry. The authors thank Ms. M. Hirahara for providing excellent technical assistance. We also thank Dr Kazuyoshi Saito at Tobata General Hospital; Dr Kentaro Hanami and Dr Shunsuke Fukuyo at Wakamatsu Hospital of the University of Occupational and Environmental Health; Dr Keisuke Nakatsuka at Fukuoka Yutaka Hospital, and all staff members at Kitakyushu General Hospital and Shimonoseki Saiseikai Hospital for their engagement in data collection of the FIRST registry.

Contributors All authors were involved in the drafting and critical revision of the manuscript, and all authors approved the final version to be published. YM had full access to all of the data in the study and YT unifies the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study conception and design: YM, KN and YT. Acquisition of data: YM, KN and YI. Analysis and interpretation of data: YM, KN, SN, SK, YI, YF and YT.

Funding This work was supported in part by a Grant-In-Aid for Scientific Research from the University of Occupational and Environmental Health, Japan, through University of Occupational & Environmental Health, Japan (UOEH) for Advanced Research (19K17919).

Competing interests KN has received speaking fees from Bristol-Myers, Sanofi, AbbVie, Eisai, Eli Lilly, Chugai, Pfizer, Takeda, and Mitsubishi-Tanabe, and research grants from Mitsubishi-Tanabe and Eisai. SN has received speaking fees from Bristol-Myers, UCB, Astellas, AbbVie, Eisai, Pfizer, Takeda and has received research grants from Mitsubishi-Tanabe, Novartis and MSD. S. Kubo has received speaking fees from Bristol-Myers. YT has received consulting fees, speaking fees, and/or honoraria from AbbVie, Daiichi-Sankyo, Chugai, Takeda, Mitsubishi-Tanabe, Bristol-Myers, Astellas, Eisai, Janssen, Pfizer, Asahi-kasei, Eli Lilly, GlaxoSmithKline, UCB, Teijin, MSD, and Santen, and received research grants from Mitsubishi-Tanabe, Takeda, Chugai, Astellas, Eisai, Taisho-Toyama, Kyowa-Kirin, AbbVie, and Bristol-Myers.

Patient consent for publication Not required.

Ethics approval The FIRST registry includes RA patients who initiated treatment with molecular-targeting rheumatic drugs in a registry study conducted by the University of Occupational and Environmental Health, Japan and cooperated by its affiliated and multiple facilities. This study was approved by the ethics review board of the University of Occupational and Environmental Health, Japan (approval number #04-23), and informed consent was obtained from all patients of the FIRST registry.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Yusuke Miyazaki <http://orcid.org/0000-0001-5692-3881>

Satoshi Kubo <http://orcid.org/0000-0001-9693-9263>

Yoshiya Tanaka <http://orcid.org/0000-0002-0807-7139>

REFERENCES

- Kvien TK. Epidemiology and burden of illness of rheumatoid arthritis. *Pharmacoeconomics* 2004;22:1–12.
- Smolen JS, Aletaha D, Barton A, et al. Rheumatoid arthritis. *Nat Rev Dis Primers* 2018;4:18001.
- Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet* 2001;358:903–11.
- Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016;388:2023–38.
- Tanaka Y. The JAK inhibitors: do they bring a paradigm shift for the management of rheumatic diseases? *Rheumatology* 2019;58:1–3.
- Clark JD, Flanagan ME, Telliez J-B. Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. *J Med Chem* 2014;57:5023–38.
- Fridman JS, Scherle PA, Collins R, et al. Selective inhibition of JAK1 and JAK2 is efficacious in rodent models of arthritis: preclinical characterization of INCB028050. *J Immunol* 2010;184:5298–307.
- Strand V, Lee EB, Fleischmann R, et al. Tofacitinib versus methotrexate in rheumatoid arthritis: patient-reported outcomes from the randomised phase III oral start trial. *RMD Open* 2016;2:e000308.
- Fleischmann R, Schiff M, van der Heijde D, et al. Baricitinib, methotrexate, or combination in patients with rheumatoid arthritis and no or limited prior disease-modifying antirheumatic drug treatment. *Arthritis Rheumatol* 2017;69:506–17.
- 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.
- 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.
- 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.
- Tanaka Y, Fautrel B, Keystone EC, et al. Clinical outcomes in patients switched from adalimumab to baricitinib due to non-response and/or study design: phase III data in patients with rheumatoid arthritis. *Ann Rheum Dis* 2019;78:890–8.
- Genovese MC, Kremer J, Zamani O, et al. Baricitinib in patients with refractory rheumatoid arthritis. *N Engl J Med* 2016;374:1243–52.
- McInnes IB, Byers NL, Higgs RE, et al. Comparison of baricitinib, upadacitinib, and tofacitinib mediated regulation of cytokine signaling in human leukocyte subpopulations. *Arthritis Res Ther* 2019;21:183.
- Kubo S, Nakayamada S, Sakata K, et al. Janus kinase inhibitor Baricitinib modulates human innate and adaptive immune system. *Front Immunol* 2018;9:1510.
- Ram N, Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. *Int J Behav Dev* 2009;33:565–76.
- Kubo S, Nakayamada S, Nakano K, et al. Comparison of efficacy of TNF inhibitors and abatacept in patients with rheumatoid arthritis; adjusted with propensity score matching. *Clin Immunol* 2018;191:67–74.
- Kawabe A, Nakano K, Kubo S, et al. Differential long-term retention of biological disease-modifying antirheumatic drugs in patients with rheumatoid arthritis by age group from the first registry. *Arthritis Res Ther* 2020;22:136.
- Ochi S, Saito K, Mizoguchi F, et al. Insensitivity versus poor response to tumour necrosis factor inhibitors in rheumatoid arthritis: a retrospective cohort study. *Arthritis Res Ther* 2020;22:41.
- 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.
- 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.
- Felson DT, Smolen JS, Wells G, et al. American College of Rheumatology/European League against rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011;63:573–86.
- Felson DT, Smolen JS, Wells G, et al. American College of Rheumatology/European League against rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;70:404–13.
- Courvoisier DS, Alpizar-Rodriguez D, Gottenberg JE, et al. Rheumatoid arthritis patients after initiation of a new biologic agent: trajectories of disease activity in a large multinational cohort study. *EBioMedicine* 2016;11:302–6.
- Lee YH, Song GG. Relative efficacy and safety of tofacitinib, baricitinib, upadacitinib, and filgotinib in comparison to adalimumab in patients with active rheumatoid arthritis. *Z Rheumatol* 2020;79:785–96.
- Fautrel B, Zhu B, Taylor PC, et al. Comparative effectiveness of improvement in pain and physical function for baricitinib versus adalimumab, tocilizumab and tofacitinib monotherapies in rheumatoid arthritis patients who are naive to treatment with biologic or conventional synthetic disease-modifying antirheumatic drugs: a matching-adjusted indirect comparison. *RMD Open* 2020;6:e001131.
- Bae S-C, Lee YH. Comparison of the efficacy and safety of tofacitinib and baricitinib in patients with active rheumatoid arthritis: a Bayesian network meta-analysis of randomized controlled trials. *Z Rheumatol* 2019;78:559–67.
- Bechman K, Subesinghe S, Norton S, et al. A systematic review and meta-analysis of infection risk with small molecule JAK inhibitors in rheumatoid arthritis. *Rheumatology* 2019;58:1755–66.