

BMJ Open Discontinuation of tofacitinib and TNF inhibitors in patients with rheumatoid arthritis: analysis of pooled data from two registries in Canada

Mohammad Movahedi ^{1,2}, Denis Choquette,³ Louis Coupal,³ Angela Cesta,¹ Xiuying Li,¹ Edward C Keystone,⁴ Claire Bombardier,^{1,2,4} OBRI Investigators¹

To cite: Movahedi M, Choquette D, Coupal L, *et al*. Discontinuation of tofacitinib and TNF inhibitors in patients with rheumatoid arthritis: analysis of pooled data from two registries in Canada. *BMJ Open* 2023;**13**:e063198. doi:10.1136/bmjopen-2022-063198

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-063198>).

Received 30 March 2022
Accepted 05 February 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Toronto General Hospital Research Institute, University Health Network, Toronto, Ontario, Canada

²HMPE, University of Toronto, Toronto, Ontario, Canada

³Department of Rheumatology, Institut de Rhumatologie de Montréal, Montreal, Québec, Canada

⁴Department of Medicine, University of Toronto, Toronto, Ontario, Canada

Correspondence to

Dr Mohammad Movahedi; mmovahed@uhnresearch.ca

ABSTRACT

Objectives The similarity in retention of tumour necrosis factor inhibitors (TNFi) and tofacitinib (TOFA) was previously reported separately by the Ontario Best Practices Research Initiative and the Quebec cohort Rhumadata. However, because of small sample sizes in each registry, we aimed to confirm the findings by repeating the analysis of discontinuation of TNFi compared with TOFA, using pooled data from both these registries.

Design Retrospective cohort study.

Setting Pooled data from two rheumatoid arthritis (RA) registries in Canada.

Participants Patients with RA starting TOFA or TNFi between June 2014 and December 2019 were included. A total of 1318 patients were included TNFi (n=825) or TOFA (n=493).

Outcome measures Time to discontinuation was assessed using Kaplan-Meier survival and Cox proportional hazards regression analysis. Propensity score (PS) stratification (deciles) and PS weighting were used to estimate treatment effects.

Results The mean disease duration in the TNFi group was shorter (8.9 years vs 13 years, $p<0.001$). Prior biological use (33.9% vs 66.9%, $p<0.001$) and clinical disease activity index (20.0 vs 22.1, $p=0.02$) were lower in the TNFi group.

Discontinuation was reported in 309 (37.5%) and 181 (36.7%) TNFi and TOFA patients, respectively. After covariate adjustment using PS, there was no statistically significant difference between the two groups in discontinuation due to any reason HR=0.96 (95% CI 0.78 to 1.19, $p=0.74$) as well as discontinuation due to ineffectiveness only HR=1.08 (95% CI 0.81 to 1.43, $p=0.61$).

TNFi users were less likely to discontinue due to adverse events (AEs) (adjusted HRs: 0.46, 95% CI 0.29 to 0.74; $p=0.001$). Results remained consistent for firstline users.

Conclusions In this pooled real-world data study, the discontinuation rates overall were similar. However, discontinuation due to AEs was higher in TOFA compared with TNFi users.

INTRODUCTION

Rheumatoid arthritis (RA) is an immune-mediated systemic inflammatory disease

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Rheumatoid arthritis population-based retrospective pooled cohort using data from two registries is the major strength of this study.
- ⇒ Using multiple imputation for missing data as well as propensity score for balancing covariates between treatment arms, we reduced the risk of bias estimate for this observational study.
- ⇒ Risk of unmeasured variables and likely systematic differences in the practice patterns of physicians participating in two registries can be considered as a limitation.

typically affecting the synovial membrane of small and large joints in a symmetrical pattern and is often associated with extra-articular manifestations.^{1 2} Left untreated, the disease frequently leads to joint damage, loss of function and quality of life. It is also associated with increased mortality secondary to different comorbidities such as cardiovascular events and infections. Agents targeting tumour necrosis factor (TNF) were the first group of agents shown to have an improved efficacy over conventional synthetic disease modifying antirheumatic drugs (csDMARDs) such as methotrexate (MTX) and hydroxychloroquine (HCQ). Tofacitinib (TOFA) is an oral, small molecule drug that can be used as an alternative to biological disease modifying antirheumatic drugs (bDMARDs) including tumour necrosis factor inhibitors (TNFis). The efficacy and safety of TOFA as monotherapy or in combination with csDMARDs has been compared with placebo in several phase III randomised clinical trials.³⁻⁶ Since TOFA approval as the first Janus kinase inhibitor for the treatment of RA, the effectiveness, safety and durability of this drug in observational studies using real-world data, has been an area of interest.⁷⁻¹⁴ The durability of a drug reflects a combination of effectiveness, safety

and tolerability and therefore a critical element in decision making regarding the utility of a drug in practice. Unfortunately, many patients are affected by nuisance adverse events that eventually lead to its discontinuation.

To date, the durability of TOFA versus TNFi has been controversial. A previous study showed that TOFA to be more durable than TNFi agents.⁸ However, recent preliminary data from several observation cohorts in contrast have shown a similar durability of TOFA compared with TNFi.^{12 15} The Ontario Best Practices Research Initiative (OBRI) and Rhumadata also previously presented independent abstracts showing similar retention of TNFi and TOFA.^{16 17} However, due to a concern about sample size and study power, particularly for the TOFA group, we proposed to evaluate the discontinuation of TNFi compared with TOFA, overall and by specific reason of discontinuation, using pooled data from both registries. We hypothesised that discontinuation of TOFA and TNFi is similar. The results of this analysis could provide us information that could help prescribers to better select the optimal usage of these agents.

METHODS

Data sources

The OBRI is a multicentre registry across Ontario, Canada, collecting data from rheumatologists and patients with RA at enrolment and follow-up. It incorporates rheumatologist assessments from approximately one-third of the rheumatologists in the province of Ontario. Patients are eligible to be enrolled if they are ≥ 16 years of age at the time of diagnosis, ≥ 18 years of age at enrolment, have a rheumatologist confirmed RA diagnosis and have at least one swollen joint. Enrolled patients are interviewed every 6 months by phone and seen by their rheumatologist in routine care. At enrolment, patients are asked for their general medical history including comorbidity status. Rheumatologists are also expected to report any history of previous comorbidity including cardiovascular disease and RA disease activity including inflammatory markers, patient global, physician global, and tender and swollen joint counts. Data on sociodemographics, smoking status, height, weight and any prior and current medications are recorded during the rheumatologist enrolment visit or the patient's interview. Patient-reported outcomes for functional status are also collected. At follow-up visits, all the information mentioned previously is updated. RA medication changes (including discontinuation and reasons for discontinuation) between visits are also captured. Rheumatologists report any incident of comorbidity and reassess disease activity during every follow-up visits.

The Rhumadata clinical database and registry monitors the clinical care of all the patients with inflammatory diseases seen at the *Institut de Rhumatologie de Montréal* (IRM), the *Centre de l'ostéoporose et de rhumatologie de Québec* (CORQ) and the *Clinique de santé Jacques-Cartier* (CSJC), the largest rheumatological clinics in

the province of Québec, Canada. Rhumadata has been collecting real-world observational data since 1998. The database currently includes the treatment history of more than 6000 patients with inflammatory disease (RA, ankylosing spondylitis, spondyloarthritis). Data collected at all visits include: demographics, disease history, laboratory values (rheumatoid factor (RF), anticirculated protein antibody, C reactive protein (CRP) and erythrocyte sedimentation rate (ESR)), all disease activity scores (disease activity score – CRP and ESR, clinical disease activity index (CDAI) and simplified disease activity index (SDAI)), patient report outcomes (PROs) including health assessment questionnaire disability index (HAQ-DI), morning stiffness, patient global evaluation of disease activity, patient evaluation of pain and physician global evaluation of disease activity. Comorbidities including, but not limited to, cardiovascular disease, diabetes, high blood pressure, cancer and infections are also collected. Medication usage for disease control are entered in the database (start and termination data as well a reason for termination).

Study population and data collection

For this study, patients with RA enrolled in the OBRI and Rhumadata registries starting their TOFA or TNFi between 1 June 2014 and 31 December 2019 (defined as end of study) were included. 1 June was chosen because the first approval date of TOFA was 14 June 2014. This allowed us to avoid selection bias as it allowed for an equal chance of choosing either TNFi or TOFA. Patients were followed from treatment start until discontinuation, death, loss to follow-up or last visit, whichever came first. The discontinuation date was defined as the first missed dose of therapy. Temporary stops of ≤ 180 days and ≤ 30 days after which the patients restarted the same medication as individual TNFi or TOFA, respectively, were counted as continuous use. We have used temporary stop definition in our previous publications from two registries.^{11 18 19}

In the OBRI, using a case report form at each visit, physicians report changes to medications and reason for any discontinuations. Reasons for discontinuation are reported as one of the following: (1) primary failure (lack of response), (2) secondary failure (failure to maintain response after 3 months loss of response), (3) adverse event, (4) reimbursement issues, (5) patient decision, (6) physician decision, (7) improvement, (8) completed treatment, (9) dosage change, (10) pregnancy and (11) other. In Rhumadata, data are collected at all visits from the respective clinics. Patients provide responses to all PRO questionnaire using an electronic touch pad or a tablet. All disease and medication changes that have occurred since the last visit are entered into the database, and the treating physician assesses the following: (1) primary failure (lack of response), (2) secondary failure (failure to maintain response after 3 months loss of response), (3) adverse event, (4) reimbursement issues, (5) patient decision, (6) physician decision, (7) improvement,

(8) completed treatment, (9) dosage change and (10) pregnancy.

For this study, dosage change was not considered a reason for discontinuation. Ineffectiveness was defined as treatment discontinuation due to primary or secondary failure.

Statistical analysis

All analyses were conducted on the primary analysis population. Descriptive statistics specifically mean and SD for continuous variables and counts and proportions for categorical variables were generated for all baseline (defined as start of medication) characteristics. Comparisons between patients on TNFi versus TOFA were conducted using the independent-samples t-test for continuous variables and the χ^2 or the Fisher's exact test for categorical variables. We first evaluated time to discontinuation due to any reason using adjusted Kaplan-Meier survival analysis (non-parametric models) and Cox proportional hazards regression (HR) analysis (parametric models) for TNFi versus TOFA users. We additionally assessed time to discontinuation due to only lack/loss of response and due to only AEs by excluding other reasons. The proportional hazards assumption was checked by including covariate interactions with time as predictors in the Cox models. Missing data in observational studies are common. This complicates data manipulation and analysis and often may produce biased and inefficient estimators. To deal with this problem, multiple imputation (imputation chained equation, $n=20$) were used to deal with missing data for covariates at treatment start. This model is commonly used under the assumption of missing at random. Twenty datasets were imputed and results were combined using Rubin's rules.^{20 21}

Estimating the propensity score (PS)

Here, treatment attribution is not random and may rely on patient specific factors. We wish to reduce the bias called 'confounding by indication'. PS matching is a method used to reduce bias in the estimation of treatment effects with observational data sets. A PS is the probability of treatment attribution based on selected patient characteristics. We estimated PS for covariates with an absolute standardised difference greater than 0.1 between the two treatment groups (table 1) to adjust for confounding by indication. We have used this approach in our previous published study comparing retention TNFi versus non-TNFi.¹⁹

PS implementation and discontinuation estimation

We compared discontinuation in TNFi versus TOFA users with Cox proportional hazard regression models. Results were presented as HRs and 95% CIs. We estimated the treatment effect using a PS stratification (deciles) approach, which has been shown to remove up to 90% of the bias in the unadjusted estimate.²² We also conducted an analysis using PS weighting including the stabilised inverse probability of treatment weight (SIPTW).

Stabilised weights were used to produce appropriate of the estimated treatment effect and avoids generating an underestimate of the variance of the estimate of the effect caused by unbalanced IPTW.²³⁻²⁵ The estimated weights were incorporated into a Cox regression that only included the treatment variables.

We combined multiple imputation with PS using a Within approach. In this approach, PS individually used to obtain treatment effect estimates in each imputation are combined to produce an overall estimate.²⁶ To create a PS weighted survival curve, we applied a SAS Macro to one imputed dataset.²⁷

PS evaluation

We used two approaches to assess the quality of the PSs estimated: (1) the goodness of fit for the multivariable logistic model was assessed using a Hosmer-Lemeshow test; (2) we also compared the distribution of PS across the treatments within strata (deciles).

Subset analysis

We conducted two subset analyses comparing discontinuation of TOFA with and without MTX and TNFi with and without MTX. Concomitant MTX use was defined as MTX use for at least 80% of the time the patient was also using TNFi or TOFA. Although arbitrary, we mimic the use of an 80% threshold to differentiate adherent from non-adherent patients, as proposed by Haynes.²⁸

Sensitivity analysis

To check the consistency of the results, we repeated the analyses in first line users of TOFA or TNFi.

All statistical analyses were conducted using SAS V.9.4 (SAS Institute).

Patient and public involvement

Patients or the public were not involved in this specific research project.

RESULTS

Patient sociodemographics and clinical characteristics

Overall study population

A total of 1318 patients (721 in the OBRI and 597 in the Rhumadata) started TNFi or TOFA on or after 1 June 2014 (figure 1), among whom 825 (62.6%) were being treated with TNFi (list of individual TNFi is provided in online supplemental table S1 and the remaining (37.4%) with TOFA. The average age (SD) of the total cohort was 57.6 (12.6) years, with the majority of patients being female ($n=1067$; 81.0%). At baseline, the mean (SD) RA duration in the total cohort was 10.5 (9.8) years, with the majority of patients being RF positive ($n=843$; 68.3%). The mean (SD) CDAI was 20.9 (12.0), and HAQ-DI was 1.2 (0.7). The mean (SD) number of comorbidities was 0.9 (1.2) for the whole cohort. Fifty-four per cent of patients had no prior exposure to bDMARDs (biological naïve). The number of patients on concurrent methotrexate (MTX) was 746 (56.6%) (table 1).

Table 1 Patient baseline characteristics: overall and by treatment group

	Primary data before multiple imputation			P value	Absolute standard difference between two treatment groups after multiple imputation	
	Total* (n=1318)	TNFi* (n=825)	TOFA* (n=493)		Before propensity score (SIPTW)	After propensity score (SIPTW)
Gender, female (%)	1067 (81.0)	650 (78.8)	417 (84.6)	0.01	0.15	0.03
Age (years), mean±SD	57.6±12.6	56.5±13.1	59.5±11.5	<0.001	0.24	0.05
Disease duration (years)	n=1315	n=824	n=491			
Mean±SD	10.5±9.8	8.9±9.3	13.0±10.1	<0.001	0.42	0.03
Positive RF	n=1234	n=775	n=459			
Positive (%)	843 (68.3)	507 (65.4)	336 (73.2)	0.01	0.16	0.05
SDAI	n=713	n=416	n=297			
Mean±SD	22.2±12.3	21.2±11.7	23.6±12.7	0.01	0.15	0.02
CDAI	n=787	n=457	n=330			
Mean±SD	20.9±12.0	20.0±11.7	22.1±12.4	0.02	0.13	0.02
HAQ-DI	n=927	n=563	n=364			
Mean±SD	1.2±0.7	1.2±0.7	1.3±0.7	0.002	0.14	0.03
ESR	n=993	n=598	n=395			
Mean±SD	21.0±19.0	20.2±17.8	22.3±20.7	0.08	0.11	0.08
Number of main comorbidity	n=664	n=373	n=291			
Mean±SD	0.9±1.2	0.8±1.1	0.9±1.2	0.54	0.07	–
Patients with no prior use of biologicals	708 (53.7)	545 (66.1)	163 (33.1)	<0.001	0.69	–
Number of prior biologicals used, mean±SD	1.38±1.54	1.03±1.19	1.97±1.84	<0.001	0.61	0.03
Patients with at least 80% time using MTX, n (%)	746 (56.6)	497 (60.2)	249 (50.5)	<0.001	0.20	0.02
Patients with at least 80% time using HCQ	446 (33.8)	312 (37.8)	134 (27.2)	<0.001	0.23	0.02
Patients with at least 80% time using LEF	177 (13.4)	92 (11.2)	85 (17.2)	<0.001	0.18	0.08
Concomitant use of steroids	396 (30.0)	226 (27.4)	170 (34.5)	0.007	0.15	0.08

Bold: statistically significant p values <0.05.

*Number of available data (N) was presented when the complete data were not available.

CDAI, clinical disease activity index; ESR, erythrocyte sedimentation rate; HAQ-DI, health assessment questionnaire disability index; HCQ, hydroxychloroquine; LEF, leflunomide; MTX, methotrexate; RF, rheumatoid factor; SDAI, simplified disease activity index; SIPTW, stabilised inverse probability of treatment weight; TNFi, tumour necrosis factor inhibitor; TOFA, tofacitinib.

TNFi versus TOFA treatment

Key differences in patient sociodemographics, clinical and medication profiles by type of therapy (TNFi vs TOFA) are summarised in [table 1](#). Patients treated with TNFi were significantly younger compared with TOFA users (mean 56.5 vs 59.5 years; $p<0.001$). Patients in the TOFA group were more likely to be female (84.6% vs 78.8%; $p=0.03$). No other significant sociodemographic differences were observed between groups.

With respect to clinical and medication profiles, compared with patients treated with TNFi, those treated with TOFA were more likely to have a longer mean disease duration (13.0 vs 8.9 years; $p<0.001$), more likely to be RF positive (73.2% vs 65.4%; $p=0.01$), less likely to

be bDMARDs naïve (33.1% vs 66.1%, $p<0.001$), less likely to use concomitant MTX (50.5% vs 60.2%; $p<0.001$) and HCQ (27.2% vs 37.8%; $p<0.001$) and more likely to use steroids (34.5% vs 27.4%; $p=0.007$) and leflunomide (LEF) (17.2% vs 11.2%; $p<0.001$) ([table 1](#)). Furthermore, at baseline, patients on TOFA had significantly worse physical function, as indicated by the higher mean HAQ-DI (1.3 vs 1.2; $p=0.002$). There was no statistically significant difference for the mean number of comorbidity between two groups (0.8 vs 0.9; $p=0.54$).

The absolute standardised difference between the two treatment groups at baseline was more than 10% for age, gender, RA duration, RF positivity, CDAI, SDAI, HAQ-DI, number of prior biologicals used, HCQ, MTX, LEF and

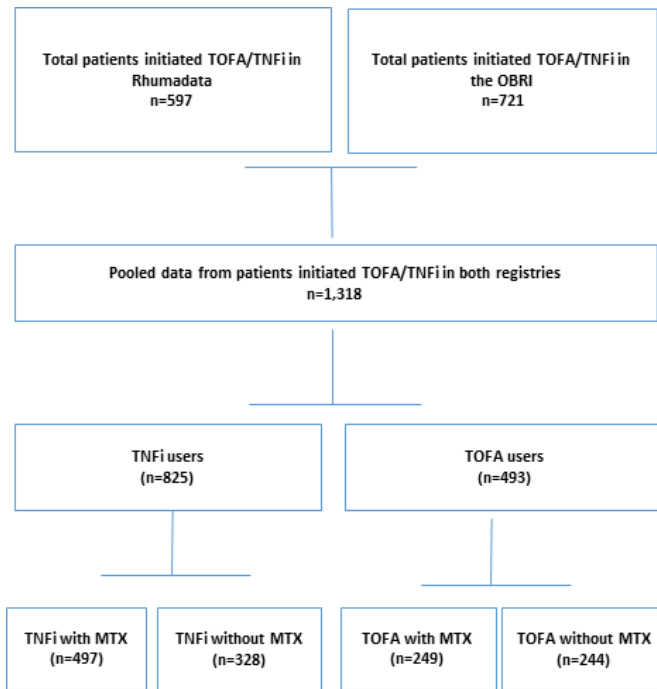


Figure 1 Study flow chart.

steroids concomitant use. Thus, these covariates were used to calculate PSs (table 1). Concomitant use was defined as using MTX, HCQ, LEF and steroids for at least 80% of the time while on TNFi or TOFA.

Evaluation of PS

The Hosmer-Lemeshow test was not significant (p=0.56) showing that the logistic regression model fits our data well for balancing covariates between the two treatment groups. The Hosmer-Lemeshow test is a statistical test for goodness of fit for logistic regression models. Like most goodness of fit tests, small p values (usually under 5%) mean that the model is not a good fit.

In addition, distribution of PSs across treatment groups within deciles was very similar (online supplemental figure S1).

Time to discontinuation by treatment

Over a mean follow-up of 23.2 months, discontinuation was reported in 309 (37.5%) and 181 (36.7%) of all TNFi and TOFA patients, respectively. The most common reason for discontinuation was ineffectiveness (54.7% (169/309) and 53.6% (97/181) for TNFi and TOFA, respectively) followed by adverse events ((12.6% (39/309) and 35.0% (46/181) for TNFi and TOFA), respectively). The median drug survival (retention) was 51.0 months (41.1-NE) and 49.4 months (37.8-NE) in TNFi and TOFA treatment groups, respectively.

In the univariable survival analysis, patients using TNFi therapy had borderline significant lower discontinuation due to any reason (unadjusted HR=0.83 (95% CI 0.69 to 1.00, p=0.05)) compared with patients using TOFA therapy (table 2). There was no significant difference for discontinuation due to ineffectiveness between the two

Table 2 Discontinuation of TNFi versus TOFA by reason for discontinuation: univariable and multivariable Cox regression models

Reason	TNFi versus TOFA		Multiple imputed data	
	Unadjusted n=1318		Propensity score stratification (PS deciles) n=1318	
	Number of events	HRs (95% CI), p value	Number of events	HRs (95% CI), p value
Due to any reason	491	0.83 (0.69–1.00), 0.05	491	0.96 (0.78–1.19), 0.74
Due to ineffectiveness	266	0.86 (0.67–1.10), 0.22	266	1.08 (0.81–1.43), 0.61
Due to adverse events	58	0.44 (0.29–0.67), 0.0002	58	0.46 (0.29–0.74), 0.001

Bold: statistically significant p value.

Covariates with more than 10% standardised difference between the two treatments groups were used to calculate propensity score: age, gender, disease duration, RF status, number of prior biologic used, HAQ-DI, ESR, CDAI, SDAI, concomitant use of MTX, LEF, HCQ and steroids. CDAI, clinical disease activity index; ESR, erythrocyte sedimentation rate; HAQ-DI, health assessment questionnaire disability index; HCO, hydroxychloroquine; LEF, leflunomide; MTX, methotrexate; PS, propensity score; RF, rheumatoid factor; SDAI, simplified disease activity index; SIPTW, stabilized inverse probability of treatment weight; TNFi, tumor necrosis factor inhibitor; TOFA, Tofacitinib.

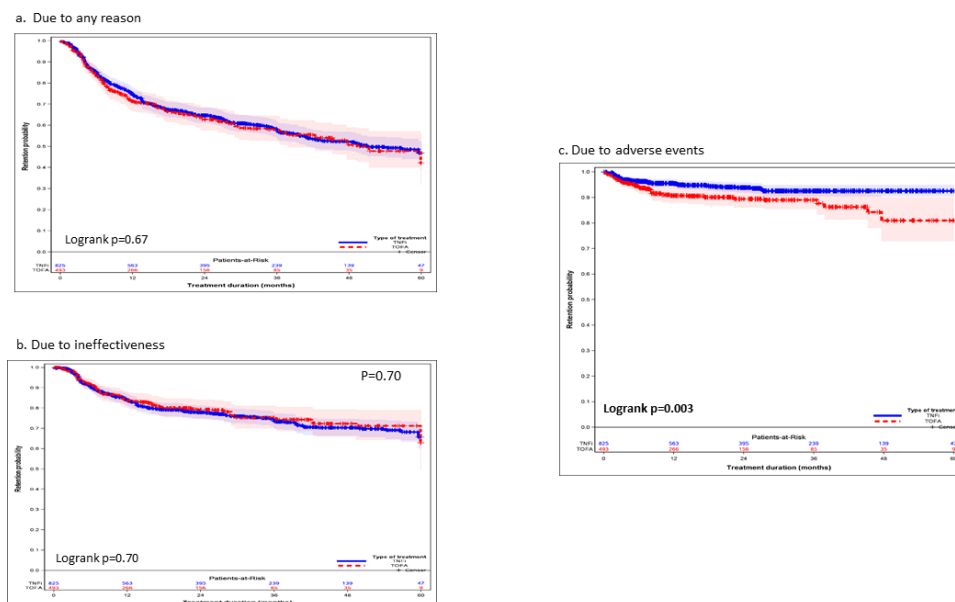


Figure 2 Propensity score weighted survival curves for time to discontinuation of TNFi versus TOFA. TNFi, tumour necrosis factor inhibitor; TOFA, tofacitinib.

groups. However, the univariable analysis showed that patients in the TNFi group were significantly less likely to discontinue their medication due to AEs compared with patients in the TOFA group (unadjusted HR=0.44 (95% CI 0.29 to 0.67, $p=0.0002$)).

Figure 2 shows the time to discontinuation by type of treatment. After using the PS weighted survival curve, no significant differences were observed in terms of discontinuation due to any reason (figure 2a; HR=0.97 (95% CI 0.79 to 1.19, $p=0.67$)), or due to ineffectiveness (figure 2b; HR=1.07 (95% CI 0.81 to 1.42, $p=0.70$)). The difference in discontinuation due to AEs between TNFi and TOFA remained significant after using the PS weighted survival curve (figure 2c; HR=0.48 (95% CI 0.30 to 0.76, $p=0.003$)).

On adjustment for stratification (deciles) and stabilised IPTW across 20 multiple imputed datasets for Cox regression models, estimated HRs for discontinuation due to any reason was not significant between the two treatment groups as shown in the table 2 (stratification: HR=0.96 (95% CI 0.78 to 1.19, $p=0.74$); stabilised IPTW: HR=0.96 (95% CI 0.79 to 1.15, $p=0.64$)).

In terms of specific reasons for discontinuation, no significant difference was observed for discontinuation due to ineffectiveness (stratification: HR=1.08 (95% CI 0.81 to 1.43, $p=0.61$); stabilised IPTW: HR=1.06 (95% CI 0.82 to 1.37, $p=0.65$)). However, compared with the TOFA treatment group, discontinuation due to AEs was statistically significantly lower in the TNFi treatment group (stratification: HR=0.46 (95% CI 0.29 to 0.74, $p=0.001$), stabilised IPTW: HR=0.51 (95% CI 0.34 to 0.77, $p=0.001$)) (table 2).

Results for first line users also remained consistent, with no differences in discontinuation of TOFA and TNFi due to any reason or ineffectiveness but higher discontinuation of TOFA due to AEs (online supplemental table S2).

Discontinuation of TOFA with and without MTX

With data restricted to TOFA users, we compared discontinuation of combination therapy (with MTX) versus monotherapy (without MTX) undertake covariate adjustment using PSs. There was no significant difference in discontinuation due to any reason (stratification: HR=1.10 (95% CI 0.81 to 1.50, $p=0.52$); stabilised IPTW: HR=0.94 (95% CI 0.71 to 1.24, $p=0.67$)), ineffectiveness (stratification: HR=1.34 (95% CI 0.87 to 2.05, $p=0.17$); stabilised IPTW: HR=0.97 (95% CI 0.66 to 1.40, $p=0.85$)), or AEs (stratification: HR=0.87 (95% CI 0.48 to 1.57, $p=0.63$); stabilised IPTW: HR=1.03 (95% CI 0.59 to 1.78, $p=0.92$)) (table 3). The results were consistent after repeating the analyses for first line users (data not shown).

Discontinuation of TNFi with and without MTX

Similar results were found for TNFi combination therapy versus TNFi monotherapy after adjusting for PSs. There was no significant difference in discontinuation due to any reason (stratification: HR=1.20 (95% CI 0.94 to 1.53, $p=0.15$); stabilised IPTW: HR=1.05 (95% CI 0.84 to 1.30, $p=0.68$)), ineffectiveness (stratification: HR=1.39 (95% CI 0.99 to 1.97, $p=0.06$); stabilised IPTW: HR=1.20 (95% CI 0.89 to 1.62, $p=0.23$)) or adverse events (stratification: HR=1.04 (95% CI 0.52 to 2.06, $p=0.91$); stabilised IPTW: HR=0.81 (95% CI 0.45 to 1.45, $p=0.47$)) (table 4). The results were consistent after repeating the analyses for firstline users (data not shown).

DISCUSSION

This real-world observational study directly compared drug discontinuation in patients with RA starting TNFi versus TOFA, overall and stratified by reason for discontinuation. Similar to other real-world studies,^{8 9 12} we found that patients starting TOFA were older, had longer

Table 3 Discontinuation of TOFA with MTX versus TOFA without MTX by reason for discontinuation: univariable and multivariable Cox regression models

		TOFA with MTX versus TOFA without MTX						
		Complete data analysis			Multiple imputed data			
Unadjusted n=493		Propensity score stratification (PS quantile) n=493			SIPTW n=1318			
Reason	Number of events	HRs (95% CI), p value	Number of events	HRs (95% CI), p value	Number of events	HRs (95% CI), p value	Number of events	
Due to any reason	181	1.05 (0.78–1.40), 0.77	71	1.25 (0.76–2.04), 0.38	181	1.10 (0.81–1.50), 0.52	181	0.94 (0.71–1.24), 0.67
Due to ineffectiveness	97	1.17 (0.79–1.75), 0.43	45	1.42 (0.76–2.67), 0.28	97	1.34 (0.87–2.05), 0.17	97	0.97 (0.66–1.40), 0.85
Due to adverse events	46	0.96 (0.54–1.72), 0.89	20	0.67 (0.26–1.75), 0.24	46	0.87 (0.48–1.57), 0.63	46	1.03 (0.59–1.78), 0.92

Covariates with more than 10% standard difference between the two treatment groups were used to calculate propensity score: ACPA status, number of prior biological used, SDAI, hypertension, CVD, cancer, concomitant use of LEF and SSZ.
The models were additionally adjusted for age and gender.
ACPA, antircululated protein antibody; CVD, cardiovascular disease; LEF, leflunomide; MTX, methotrexate; PS, propensity score; SDAI, simplified disease activity index; SIPTW, stabilised inverse probability of treatment weight; SSZ, sulfasalazine; TOFA, tofacitinib.

Table 4 Discontinuation of TNFi with and without MTX by reason for discontinuation: univariable and multivariable Cox regression models

		TNFi with MTX versus TNFi without MTX						
		Complete data analysis			Multiple imputed data			
Unadjusted n=825		PS stratification (PS quantile) n=825			SIPTW n=825			
Reason	Number of events	HRs (95% CI), p value	Number of events	HRs (95% CI), p value	Number of events	HRs (95% CI), p value	Number of events	
Due to any reason	309	1.11 (0.88–1.40), 0.38	161	1.15 (0.82–1.63), 0.42	309	1.20 (0.94–1.53), 0.15	309	1.05 (0.84–1.30), 0.68
Due to ineffectiveness	169	1.29 (0.94–1.78), 0.11	87	1.15 (0.72–1.85), 0.56	169	1.39 (0.99–1.97), 0.06	169	1.20 (0.89–1.62), 0.23
Due to adverse events	41	0.93 (0.49–1.76), 0.83	20	0.99 (0.37–2.67), 0.99	41	1.04 (0.52–2.06), 0.91	41	0.81 (0.45–1.45), 0.47

Covariates with more than 10% standard difference between the two treatment groups were used to calculate PS: education, hypertension, lung disease, ESR, CRP, fatigue, concomitant use of LEF, HCQ, SSZ and NSAIDs.
The models were additionally adjusted for age and gender.
CRP, C reactive protein; ESR, erythrocyte sedimentation rate; HCQ, hydroxychloroquine; LEF, leflunomide; MTX, methotrexate; NSAIDs, non-steroid anti-inflammatory drugs; PS, propensity score; SIPTW, stabilised inverse probability of treatment weight; SSZ, sulfasalazine; TNFi, tumour necrosis factor inhibitor.



disease duration, had a higher proportion of prior biologic use, had higher disease activity and were more commonly using it as monotherapy when compared with the TNFi group.

The most common reason for discontinuation in our study was ineffectiveness in both the TNFi (54.7%) and TOFA (53.6%) groups. Finckh *et al*⁸ in 2020 reported a similar proportion of discontinuations for these two treatment groups. In addition, Bechman *et al*²⁹ in 2020 reported 52% of patients discontinued their first TNFi, 40% due to ineffectiveness and 41% due to adverse events.

The results of our study suggest that after adjusting for differences in baseline characteristics (through PSs), discontinuation for all reasons is similar in patients with RA starting TNFi agents and those starting TOFA. Harnett *et al*¹² in 2016, in a real-world retrospective study, also found comparable persistence for TOFA treatment compared with adalimumab, abatacept and etanercept treatments. A recent study from Australia using real-world data from 1950 RA-matched patients also showed that the persistence and effectiveness of TOFA is similar to biological agents.¹⁵ Similar to our study, the median treatment persistence was comparable for TOFA and biological agents (34.2 vs 33.8 months). However, the median drug survival was higher in Canada compared with Australia (50 months vs 34 months). Similar crude discontinuation was found in TOFA and TNFi groups in a study conducted using the Swiss Clinical Quality Management registry that included almost 2000 patients starting TOFA, TNFi or non-TNFi. However, after adjusting for some potential confounding factors, they found a higher risk of discontinuation due to any reason (HR 1.29 (95% CI 1.14 to 1.47)) and due to ineffectiveness ((HR: 1.59 (95% CI 1.33 to 1.89)) in TNFi users compared with TOFA users.⁸

We found the discontinuation of TNFi due to adverse events was lower when compared with TOFA. Similar to our findings, Finckh *et al*⁸ in 2020 found lower discontinuation due to adverse events in TNFi compared with TOFA users (adjusted HR: 0.76, 95% CI 0.59 to 0.98).

ORAL Surveillance study (open label) also showed that risk of major adverse cardiovascular events and malignancy was higher in the combined TOFA doses (3.4% and 4.2%, respectively) compared with TNFi users (2.5% and 2.9%, respectively) in patients with RA over 50 years old.³⁰

Other real-world studies have shown a higher risk of severe infection and herpes zoster for TOFA users compared with other biologics. Pawar *et al*, in 2020, in a multidatabase cohort study, found that the risk of hospitalisations due to serious infection in patients with RA using TOFA was higher (propensity adjusted HR: 1.41 (95% CI 1.15 to 1.73)) compared with etanercept and lower compared with infliximab (propensity adjusted HR (0.81, 0.65–1.00)).¹⁴ Curtis *et al*¹³ in 2016, in a real-world study, showed that after multivariable adjustment, the risk for HR associated with TOFA was 2.01 (95% CI 1.40 to 2.88) compared with abatacept.

The percentage of patients in our study who discontinued TOFA due to adverse events was 9.3% (46/493), which is similar to that reported in most of the literature (3.2%–9.6%).^{10 31–35} However, some studies have shown a higher percentage of discontinuation due to adverse events.^{36 37} For example, Pope *et al* in 2019, in a long-term extended study found 25.7% of patients discontinued TOFA due to any adverse events and 17.8% discontinuation due to drug-related adverse events.³⁷

One possible explanation for the higher discontinuation due to AEs for TOFA (9.3%: 46/493) compared with TNFi users (4.7%: 39/825) could be the much higher use of steroids found in the TOFA group who stopped their medication due to adverse events (22/46: 48% in TOFA vs 11/39: 28% in TNFi users). However, our propensity-adjusted models comparing TNFi versus TOFA discontinuation would have taken into account this difference in steroid use between treatment groups.

Similar to our study, two observational studies reported that most of the TOFA users had previously used biological agents and 3.5%–6.0% of them discontinued them due to adverse events.^{10 32} Regarding the difference in discontinuation between monotherapy and combination therapy (with and without MTX), we found similar retention for both TOFA user groups. Other studies have also found no significant differences in improvement of disease activity when comparing TOFA used with and without MTX.^{9 10}

In this pooled analysis, we found no significant difference in discontinuation between TNFi with and without MTX. In a previous study, we showed a non-significant trend towards lower discontinuation for any reason for the csDMARD/bDMARD group compared with bDMARD monotherapy group (HR 0.76, 95% CI 0.55 to 1.05).¹⁸ Other real-world studies have shown contrasting results. Finckh *et al*⁸ showed that TNFi maintenance was decreased when prescribed without concomitant csDMARDs (HR: 1.27 (95% CI 1.08 to 1.49)). Soliman *et al* in 2011 found lower discontinuation for TNFi+MTX compared with TNFi monotherapy.³⁸ A new study from the British Society for Rheumatology Biologic Register also showed that patients receiving TNFi monotherapy were more likely to discontinue TNFi compared with patients receiving TNFi with MTX.²⁹ However, their findings were only significant for the ≤75 years old cohort but not for the older cohort. Reed *et al* in 2019 showed that the CDAI LDA/remission was significantly higher, and mean pain score at 6 months was significantly lower, for patients receiving TNFi combination therapy compared with TNFi monotherapy.⁹ However, they did not find a significant difference among patients who started TNFi as a fourth-line treatment.⁹

One possible explanation for these inconsistent results from studies on TNFi mono and combination therapy may be the difference in the definition used for combination MTX therapy. While we defined concomitant use of MTX as using MTX for at least 80% of the time while receiving TNFi, there was no clear definition of

concomitant use in the other studies. Another possible explanation is that one-third of the TNFi patients in our study were etanercept users, which have been shown to have the same risk of discontinuation when monotherapy and combination therapy are compared.³⁹ Favalli *et al*,⁴⁰ in 2016, in a long-term treatment study, also showed that the median survival for both etanercept monotherapy and combination therapy were similar (>53.5 months).

Strengths of our study include using multicentre data, controlling for disease severity, comorbidities and demographics to balance measured prognostic factors by adjusting our models for PSs (stratification and SIPTW).

There are some limitations of this study. Given the lack of randomisation, the estimates might be biased as some patient and physician variables were unmeasured. Additionally, there are likely systematic differences in the practice patterns of physicians participating in the OBRI and Rhumadata. There may have been differences in the number of follow-up visits from both registries, but we did include registry as a covariate. In terms of reported adverse events, our data limited to difference in the granularity of safety data from both registries, precluding an accurate description. We also did not investigate the discontinuation due to other reason in this study.

In summary, this study demonstrates that discontinuation of therapy is similar in patients started on TNFi and TOFA therapies. There was no significant difference in discontinuation between groups due to ineffectiveness. However, we found a significant difference in discontinuation due to adverse events. Further analyses are still needed to complement this information.

Acknowledgements We would like to thank all Ontario Best Practices Research Initiative (OBRI) and Rhumadata participants and following investigators

Contributors All authors contributed to the conception or design of the work. MM conducted the data analysis and drafted the manuscript. DC, CB, LC, EK, AC, XL and MM revised the work critically and approved the final version of the manuscript. DC and CB contributed equally to this paper. DC and CB acted as guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests MM is employee at OBRI and has a faculty position (status) at University of Toronto with no conflict of interest; AC and XL are employees at OBRI with no conflict of interest. DC is the owner of the Rhumadata software and is the scientific director of Conceptmed-Rhumadata; LC, biostatistician at the Institut de Rhumatologie de Montréal, declares no conflict of interest; ECK has received sources of funding for research from Amgen, Merck and Pfizer; has a consulting agreement/is a member of an advisory board for AbbVie, Amgen, Celltrion, Myriad Autoimmune, F. Hoffmann-La Roche Inc, Janssen Inc, Lilly Pharmaceuticals, Merck, Pfizer Pharmaceuticals, Sandoz, Sanofi-Genzyme, Samsung Bioepis; and received speaker honoraria agreements for Amgen, AbbVie, Celltrion, F. Hoffmann-La Roche Inc, Janssen Inc., Merck, Pfizer Pharmaceuticals, Sandoz and Sanofi Genzyme. CB held a Canada Research Chair in Knowledge Transfer for Musculoskeletal Care and has received sources of funding for research for a longitudinal cohort study (Ontario Best Practices Research Initiative) including Psoriatic Arthritis cohort (PSA) from Abbvie, Amgen, Celgene, Gilead, GSK, Janssen, Lilly, Merck, Medexus, Medreleaf/Aurora, Novartis, Pfizer (Hospira), Roche, Sandoz, Sanofi/ Genzyme and UCB Canada Inc.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Institutional ethics approval was obtained for both OBRI (University Health Network research ethics board # 07–0729 AE) and Rhumadata

(Institutional Review Board Services #IRB00005290), and all patients provided informed consent before study enrolment. This study was conducted in compliance with the principles of the Declaration of Helsinki. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

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 iD

Mohammad Movahedi <http://orcid.org/0000-0001-6154-0121>

REFERENCES

- Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
- Widdifield J, Paterson JM, Bernatsky S, *et al*. The epidemiology of rheumatoid arthritis in Ontario, Canada. *Arthritis Rheumatol* 2014;66:786–93.
- 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.
- 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.
- 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.
- Strand V, Burmester GR, Zerbini CAF, *et al*. Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: patient-reported outcomes from a phase III trial. *Arthritis Care Res (Hoboken)* 2015;67:475–83.
- Caporali R, Zavaglia D. Real-World experience with tofacitinib for the treatment of rheumatoid arthritis. *Clin Exp Rheumatol* 2019;37:485–95.
- Finckh A, Tellenbach C, Herzog L, *et al*. Comparative effectiveness of antitumor necrosis factor agents, biologics with an alternative mode of action and tofacitinib in an observational cohort of patients with rheumatoid arthritis in Switzerland. *RMD Open* 2020;6:e001174.
- Reed GW, Gerber RA, Shan Y, *et al*. Real-World comparative effectiveness of tofacitinib and tumor necrosis factor inhibitors as monotherapy and combination therapy for treatment of rheumatoid arthritis. *Rheumatol Ther* 2019;6:573–86.
- Iwamoto N, Tsuji S, Takatani A, *et al*. Efficacy and safety at 24 weeks of daily clinical use of tofacitinib in patients with rheumatoid arthritis. *PLoS One* 2017;12:e0177057.
- Curtis JR, Schulze-Koops H, Takiya L, *et al*. Efficacy and safety of tofacitinib in older and younger patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2017;35:390–400.
- Harnett J, Gerber R, Gruben D, *et al*. Evaluation of real-world experience with tofacitinib compared with adalimumab, etanercept, and abatacept in RA patients with 1 previous biologic DMARD: data from a U.S. administrative claims database. *J Manag Care Spec Pharm* 2016;22:1457–71.
- Curtis JR, Xie F, Yun H, *et al*. Real-World comparative risks of herpes virus infections in tofacitinib and biologic-treated patients with rheumatoid arthritis. *Ann Rheum Dis* 2016;75:1843–7.

- 14 Pawar A, Desai RJ, Gautam N, *et al.* Risk of admission to hospital for serious infection after initiating tofacitinib versus biologic dmards in patients with rheumatoid arthritis: a multidatabase cohort study. *The Lancet Rheumatology* 2020;2:e84–98.
- 15 Bird P, Littlejohn G, Butcher B, *et al.* Real-World evaluation of effectiveness, persistence, and usage patterns of tofacitinib in treatment of rheumatoid arthritis in Australia. *Clin Rheumatol* 2020;39:2545–51.
- 16 Choquette D, Bessette L, Sauvageau LC, *et al.* SATO120 COMPARATIVE effectiveness of tofacitinib and TNF inhibitors since 2014. impact of combination with methotrexate. Annual European Congress of Rheumatology, EULAR 2019, Madrid, 12–15 June 2019; June 2019
- 17 Movahedi M, Cesta A, Li X, *et al.* Discontinuation rate of tofacitinib is similar when compared to TNF inhibitors in rheumatoid arthritis patients: real world results from a rheumatoid arthritis cohort. *ACR* 2020.
- 18 Lau AN, Thorne JC, Movahedi M, *et al.* Effect of concomitant disease-modifying antirheumatic drugs and methotrexate administration route on biologic treatment durability in rheumatoid arthritis: OBRI cohort results. *J Rheumatol* 2019;46:874–86.
- 19 Movahedi M, Hepworth E, Mirza R, *et al.* Discontinuation of biologic therapy due to lack/loss of response and adverse events is similar between tnfi and non-tnfi class: results from a real-world rheumatoid arthritis cohort. *Semin Arthritis Rheum* 2020;50:915–22.
- 20 Rubin DB. Multiple imputation for nonresponse in surveys. In: *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ, USA: John Wiley & Sons, 9 June 1987.
- 21 van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Statist Med* 1999;18:681–94.
- 22 D'Agostino RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17:2265–81.
- 23 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.
- 24 Pezzi A, Cavo M, Biggeri A, *et al.* Inverse probability weighting to estimate causal effect of a singular phase in a multiphase randomized clinical trial for multiple myeloma. *BMC Med Res Methodol* 2016;16:150.
- 25 Xu S, Ross C, Raebel MA, *et al.* Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health* 2010;13:273–7.
- 26 Granger E, Sergeant JC, Lunt M. Avoiding pitfalls when combining multiple imputation and propensity scores. *Stat Med* 2019;38:5120–32.
- 27 Meyers JP. The %newsurv family of macros: an update on the survival plotting macro %newsurv and an introduction to expansion macros. Proceeding of Pharmaceutical SAS Users Group Conference (PharmaSUG 2017) Paper DV15; Carry, NC: SAS Institute Inc, 2017
- 28 Haynes RB. *A critical review of the "determinants" of patient compliance with therapeutic regimens*. Baltimore, MD: Johns Hopkins University Press, 1976.
- 29 Bechman K, Oke A, Yates M, *et al.* Is background methotrexate advantageous in extending TNF inhibitor drug survival in elderly patients with rheumatoid arthritis? an analysis of the British Society for rheumatology biologics register. *Rheumatology (Oxford)* 2020;59:2563–71.
- 30 Ytterberg SR, Bhatt DL, Mikuls TR, *et al.* Cardiovascular and cancer risk with tofacitinib in rheumatoid arthritis. *N Engl J Med* 2022;386:316–26.
- 31 Mimori T, Harigai M, Atsumi T, *et al.* AB0431 Post-marketing surveillance of tofacitinib in japanese patients with rheumatoid arthritis: an interim report of safety data. Annual European Congress of Rheumatology, 14–17 June, 2017; June 2017:1200
- 32 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.
- 33 Sansinanea P, Costi AC, Vulcano A, *et al.* AB0434 Tofacitinib in rheumatoid arthritis: real life experience. Annual European Congress of Rheumatology, 14–17 June, 2017; June 2017:1202
- 34 Schneeberger E, Salas A, Medina L, *et al.* AB0419 Real world use of tofacitinib in rheumatoid arthritis: data from latin america. Abstract. Annual European Congress of Rheumatology, 14–17 June, 2017; June 2017
- 35 Miyazaki Y, Nakano K, Nakayamada S, *et al.* 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. *Ann Rheum Dis* 2021;80:1130–6.
- 36 Mueller R, Popp F, Mattow F, *et al.* AB0497 Effectiveness, tolerability, and safety of tofacitinib in rheumatoid arthritis: A retrospective analysis of real-world data from the st. gallen and aarau cohort. *Annual European Congress of Rheumatology, EULAR 2018, Amsterdam, 13–16 June 2018* 2018:69.
- 37 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.
- 38 Soliman MM, Ashcroft DM, Watson KD, *et al.* Impact of concomitant use of dmards on the persistence with anti-TNF therapies in patients with rheumatoid arthritis: results from the British Society for rheumatology biologics register. *Ann Rheum Dis* 2011;70:583–9.
- 39 Sebastiani M, Manfredi A, Iannone F, *et al.* Factors predicting early failure of etanercept in rheumatoid arthritis: an analysis from the Gruppo Italiano di Studio sulla early arthritis (Italian group for the study of early arthritis) registry. *Arch Rheumatol* 2020;35:163–9.
- 40 Favalli EG, Pagnola F, Biggioggero M, *et al.* Twelve-year retention rate of first-line tumor necrosis factor inhibitors in rheumatoid arthritis: real-life data from a local registry. *Arthritis Care Res (Hoboken)* 2016;68:432–9.