Comparative Effectiveness of Anti–Tumour Necrosis Factor agents, Biologics with an Alternative Mode of Action and Tofacitinib in an observational cohort of Patients with Rheumatoid Arthritis in Switzerland.

Online Supplement

Methods

Study outcome

Drug maintenance was defined as the time between the initiation (first dose) and the discontinuation of the treatment. The discontinuation was defined as the last dose plus one dispensation interval. When treatment discontinuation had not been explicitly provided by the treating rheumatologist, we used a heuristic to define a date of treatment stop[1]. In detail, treatment with a bDMARD was considered as stopped at either the date of initiation of a new bDMARD, the date of initiation of a new triple therapy of csDMARDs, or the date of death, whatever came first.[1] For treatments without an observed stop, drug maintenance was censored at the last database entry (last visit, last medication entry or last adverse event report) of the patient.

Variable definition

We here describe the covariates and specific aspects of data preparation for each one of them.

- DAS28 at baseline: represents the DAS28-crp based on the 28 swollen and tender joint count and the C-reactive protein level. If the DAS28-crp was missing, the DAS28-esr was used instead. Only data between minus three months of the treatment start and treatment start was considered. If multiple measurements were available, a weighted average weighted by the time to treatment start was used.
- Disease duration at baseline : time between diagnosis and time of treatment initiation in years

- Seropositivity: was positive if the Rheumatoid Factor and/or anti-CCP was positive. Negative if both were negative, unknown otherwise. Data for RF and anti-CCP are collected in a non-time-varying manner in the SCQM.
- Smoking at baseline: Based on patient reported data categorized as never, past or current. If a patient reported to have 'never' smoked after having reported 'past' or 'current' smoking behavior, we assumed the most recent information was valid and replaced the values by never. For baseline smoking, the closest information within five year of treatment start was assumed representative for the smoking status at treatment start.
- Body mass index (BMI) at baseline: Based on weighted average of height and weight within five years of treatment start.

Statistical analysis

All data preparation and analyses was performed using R statistical software version 3.6.1.

Analysis of drug maintenance:

The main analyses were performed after multiple imputation of missing baseline covariates (see below). The Cox proportional hazards assumption was evaluated using one of the 75 data-sets completed by multiple imputation. Schoenfeld residuals were plotted by each covariate and deviations from the proportional hazards assumption were tested using the *cox.zph* function from the R package *survival*.

To plot the predicted curves based on the multiple adjusted Cox-PH models of datasets completed with multiple imputation of missing baseline covariates, we first chose the parameter settings for a "typical" patient of this population by taking the mean for continuous

covariates and the majority category for the categorical variables sex (female), smoking (no) and seropositivity (yes). We set the number of prior b/tsDMARDs to one. We then predicted the survival rate for each time-point for this parameter setting for each dataset, derived the mean of all predictions for each time-point and plotted the means over time.

Competing risk analysis:

In the SCQM, multiple reasons can be given for the discontinuation of a drug (insufficient effectiveness, tolerance issues, remission or other). In case multiple reasons were given, we applied the following priority rule: intolerance over insufficient effectiveness over other. The above rule resulted in the three mutually exclusive discontinuation reasons: intolerance issues, insufficient effectiveness and other reasons.

In order to relate the potential effect of the b/tsDMARD treatment groups of interest on the risks of discontinuation for intolerance or insufficient effectiveness, we modelled the multiple covariate adjusted cumulative incidence for these two events separately. We used the function crr (competing risk regression) from the R package 'survival' to model the adjusted cumulative incidence function and then generated predictions for the 'typical' patient as described above for the overall retention analysis for any discontinuation reason. We then derived the survival rate from the cumulative incidence function (as 1 – cumulative incidence function) and plotted the mean of all of predictions of all datasets based on multiple imputation for missing covariates over time.

Response analysis:

Response rate (CDAI low disease activity state) at one year was analyzed for the subset of treatment courses that had been started more than one year of the database closure. Only therapies that were still under observation at 12 months (i.e. last SCQM database entry more

than 12 months after start of treatment) were considered. For treatment courses with a registered SCQM visit between 9 and 15 months, the observed CDAI value was used. For treatment courses without a follow-up visit between 9 and 15 months, a multiple adjusted linear mixed effects model was used to predict the outcome (CDAI) at 12 months. The same covariates as used in the main retention time model were used. Multiple imputation was used to account for missing baseline covariate data.

If the thus derived (predicted or observed) CDAI value at 12 months was lower ≤ 10 **and** the therapy was still ongoing after 1 year, the patient was considered to be in low disease activity state. Otherwise, if the therapy duration was shorter than 12 months **or** if the CDAI at 12 months was above 10, the patient was considered a non-responder. Logistic regression models were used to evaluate the influence of the therapies on the response.

Multiple imputation of missing baseline covariates:

To account for missing baseline covariate values, multiple imputation by chained equations was used. Full details about the variables included in the imputations are provided in **Table S1.** Since 62% of patients had one or more missing covariates, 75 imputations with 30 iterations were done for each dataset. Predictive mean matching (pmm) was used to impute continuous variables and logistic regression (logreg) for binary variables. Convergence of imputations was assessed by visual inspection of the mean and variance changes by iteration and dataset. The same models were also fitted using the subset population with complete datasets to assess the robustness of the results. Pooling of model estimates was performed according to Rubin's rules. For the imputation we used the mice package version 2.46.0 in the R statistical software (R Development Core Team, 2011) version 3.5.0.

Supplementary Table S1. Variables used in the multiple imputation models.

Variable	Used in imputation*	Comments	Predicted	Used as predictor	Method	Missing-ness	Levels
time	retention time, competing risk	Drug retention time	no	yes	pmm	0%	
censored	retention time	Censoring indicator	no	yes	pmm	0%	
failure	competing risk	Type of event	no	yes	pmm	0%	adverse event, ineffectiveness, other
log_time_rt	response analysis	log of time point of follow-up visit	no	yes	pmm	0%	
bdmard.type	all	Treatment group of interest	no	yes	polyreg	0%	Tofa, OMA, TNF
gender	all		no	yes	logreg	0%	female, male
bl.age	all		no	yes	pmm	0%	
bl.disease.duration	all		yes	yes	pmm	3%	
m.start_date	all	Start date of treatment course	no	yes	pmm	0%	
v.inclusion_date	all	First visit of patient in SCQM	no	yes	pmm	0%	
max_education_ever	all		yes	yes	polyreg	19%	compulsary, vocational, higher
bl.smoker	all		yes	yes	logreg	10%	former/never, current
previous.bdmards_cat	all	Number of previous b/ts DMARDs	no	yes	polyreg	0%	0, 1, 2 or 3+
bl.combi_type	all		no	yes	polyreg	0%	none, MTX, other
bl.swollen_joints	all		yes	yes	pmm	51%	
bl.painful_joints	all		yes	yes	pmm	52%	
bl.physician_global	all		yes	yes	pmm	54%	
bl.patient_global	all		yes	yes	pmm	67%	
bl.das28	all		yes	yes	pmm	54%	
bl.cdai	all		yes	no	Passive imputation	70%	
anti_ccp	all	Excluded from predictors in model for seropositivity	yes	yes	logreg	5%	
rheumatoid_factor	all		yes	yes	logreg	2%	
seropositivity	all	Excluded from models for RF and anti-CCP	yes	yes	Passive imputation	3%	
bl.height_m	all	Excluded from predictors in model for BMI	yes	yes	pmm	11%	
bl.weight_kg	all		yes	yes	pmm	9%	
bl.bmi	all	Excluded from models for height and weight	yes	yes	Passive imputation	12%	
bdmard.type.1.bl.combi_type.1	all	interaction btw bdmard type and mono/combo treatment	no	no			
bdmard.type.1.bl.combi_type.2	all		no	no			
bdmard.type.2.bl.combi_type.1	all		no	no			
bdmard.type.2.bl.combi_type.2	all		no	no			

Supplementary Table 2: Number of patients and treatment courses at the different steps of data preparation.

	patients	TCs	TCs_Tofa	TCs_OMA	TCs_TNF
1: all treatment courses	9616	38416	877	4444	8507
2: all, same sequential drugs merged	9616	38275	858	4374	8456
3: exclude RITUXIMAB TCs	6384	12381	854	3081	8446
4: only TCs in study period	2720	4274	854	1428	1992
5: exclude lost to follow-up immediately (*dataset for retention analysis*)	2548	3979	793	1339	1847
6: TCs started more than one year ago	2457	3766	758	1274	1734
7: TCs followed up for more than one year (*dataset for response analysis*)	2089	3122	635	1032	1455

Legend: TC: Treatment course, Tofa: Tofacitinib, OMA: other mode of action bDMARD, TNF: Tumor necrosis alpha inhibitor.

Supplementary Table 3: Complete output of Cox proportional hazard model focusing on analysing potential differences in drug maintenance between b/tsDMARD groups

	HR	2.5%-CI	97.5%-CI	P value
OMA versus Tofa*	1.09	0.96	1.24	0.2
TNF versus Tofa*	1.29	1.14	1.47	< 0.0001
Male versus female	0.98	0.87	1.06	0.39
baseline age (per year increase)	0.998	0.995	1.002	0.33
baseline disease duration (per year increase)	0.99	0.985	0.995	< 0.0001
Seroposivity yes versus no	0.97	0.88	1.06	0.48
baseline BMI (per unit increase)	1.012	1.004	1.02	0.0047
Non-smoker versus current smoker	0.98	0.88	1.09	0.74
baseline DAS28 (per unit increase)	1.02	0.97	1.08	0.46
Number of previous b/tsDMARDs: 1 versus 0	1.27	1.13	1.41	< 0.0001
Number of previous b/tsDMARDs: 2 versus 0	1.38	1.21	1.58	< 0.0001
Number of previous b/tsDMARDs: 3+ versus 0	1.55	1.36	1.76	< 0.0001
Tofa started after launch second JAK versus before	1.74	1.31	2.31	0.00015
OMA started after launch second JAK versus Tofa started before	0.76	0.53	1.10	0.15
TNF started after launch second JAK versus Tofa started before	0.80	0.57	1.12	0.19

Legend: b/tsDMARDs: Tofa reference level, bDMARD (-OMA) and TNFi. Multiple imputation was used for missing baseline covariates. Analysis based on data of 2600 patients with 4023 treatment courses and a total of 2103 discontinuation events. HR: Hazard Ratio, 2.5%-CI, 97.5%-CI: lower and upper limit of the 95% confidence interval of the hazard ratio respectively. *Reference level: Tofa treatment before start of second JAK inhibitor.

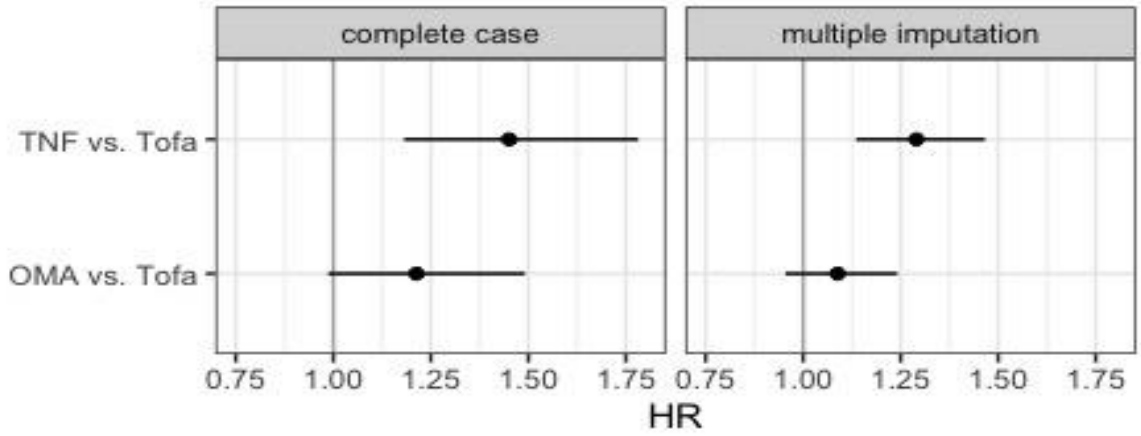
Supplementary Table S4: Results of the Cox PH model focusing on differences in b/tsDMARD monotherapy (MONO) versus in combination with csDMARDs (COMBO)

	HR	2.5%-CI	97.5%-CI	p value
OMA versus Tofa*	1.12	0.95	1.33	0.17
TNF versus Tofa*	1.27	1.08	1.49	0.0035
Tofa in mono versus combo	1.11	0.91	1.35	0.3
OMA started after launch second JAK versus Tofa started before	0.78	0.54	1.13	0.18
TNF started after launch second JAK versus Tofa started before	0.80	0.57	1.12	0.2
Male versus female	0.96	0.87	1.06	0.41
baseline age (per year increase)	0.999	0.995	1.002	0.39
baseline disease duration (per year increase)	0.989	0.984	0.994	< 0.001
Seroposivity yes versus no	0.97	0.88	1.07	0.53
baseline BMI (per unit increase)	1.01	1.01	1.02	0.002
Non-smoker versus current smoker	0.97	0.87	1.08	0.6
baseline DAS28 (per unit increase)	1.02	0.96	1.07	0.56
Number of previous b/tsDMARDs: 1 versus 0	1.25	1.12	1.39	< 0.001
Number of previous b/tsDMARDs: 2 versus 0	1.36	1.19	1.55	< 0.001
Number of previous b/tsDMARDs: 3+ versus 0	1.53	1.35	1.74	< 0.001
Tofa after launch second JAK versus before	1.71	1.28	2.28	0.0003
OMA in mono versus Tofa in combo	0.93	0.73	1.19	0.56
TNF in mono versus Tofa in combo	1.11	0.88	1.40	0.38

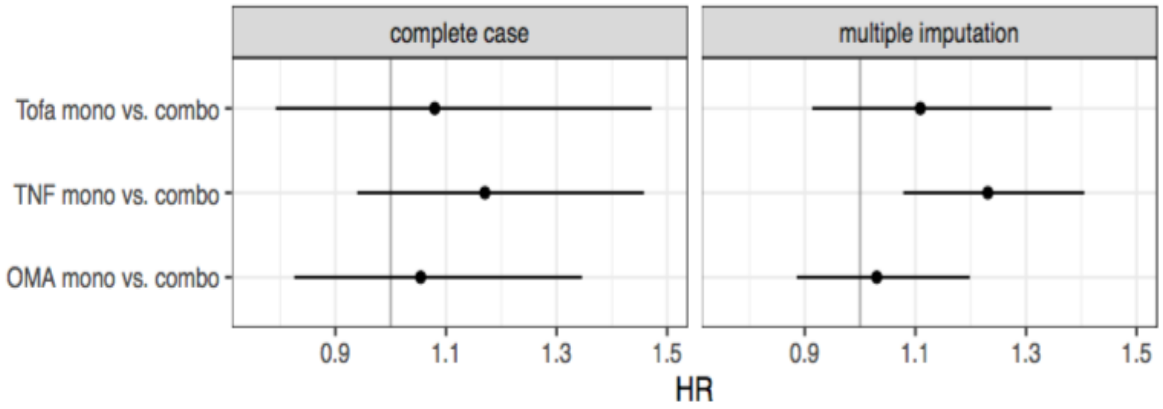
Legend: COMBO treatment and drug type-Tofacitinib are the reference level. Multiple imputation was used for missing baseline covariates. Analysis based on data of 2600 patients with 4023 treatment courses and a total of 2103 discontinuation events. HR: Hazard Ratio, 2.5%-CI, 97.5%-CI: lower and upper limit of the 95% confidence interval of the hazard ratio respectively.

Supplementary Figure S1: Treatment comparison in complete case versus after multiple imputation

A:



B:



Legend: Panel A: Summary of hazard ratios for treatment comparisons between TNFi and bDMARD-OMA and Tofa in complete case multiple adjusted models. (1515 TCs with 868 events) and models based on multiple adjusted models after imputation of baseline covariates (4023 TCs and 2103 events). **Panel B:** Summary of hazard ratios for MONO versus COMBO treatment in complete case multiple adjusted models (1515 TCs with 868 events) and models based on multiple adjusted models after imputation of baseline covariates (3978 TCs with 2199 events).

1. *Henry J, Gottenberg JE, Rouanet S, et al. Doses of rituximab for retreatment in rheumatoid arthritis: influence on maintenance and risk of serious infection. Rheumatology (Oxford) 2018;**57**(3):538-47 doi: 10.1093/rheumatology/kex446[published Online First: Epub Date]|.*