The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients

Mulligen, E. van; Weel, A.E.A.M.; Kuijper, T.M.; Hazes, J.M.W.; Helm-van Mil, A.H.M. van der; Jong, P.H.P. de

Seminars in Arthritis and Rheumatism 50 (2020) 423–431

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients

E. van Mulligen[a,*], A.E.A.M. Weel[a,b], T.M. Kuijper[b], J.M.W. Hazes[a], A.H.M. van der Helm- van Mil[a,c], P.H.P. de Jong[a]

[a] Department of Rheumatology, Erasmus MC, Rotterdam, the Netherlands
[b] Department of Rheumatology, Maasstad Hospital, Rotterdam, the Netherlands
[c] Department of Rheumatology, LUMC, Leiden, the Netherlands

ARTICLE INFO

Keywords:
Tapering
DMARDs
Rheumatoid arthritis
Patient reported outcome measures

ABSTRACT

Objectives: To determine the impact of a disease flare on patient reported outcome measures (PROMs) in rheumatoid arthritis (RA) patients, who are tapering treatment.
Methods: Data were used from the TARA trial; a multicenter, randomized controlled trial in which RA patients, with a well-controlled disease (DAS≤2.4 and SJC≤1) for at least 6 months, gradually tapered their DMARDs. PROMs of patients with a flare (DAS>2.4 and/or SJC>1) were compared every three months before and after a flare with their own norm values. Linear Mixed Models were used to investigate whether a disease flare influenced functional ability (HAQ-DI), fatigue (BRAF-MDQ), quality of life (EQ-5D and SF36), anxiety and depression (HADS), morning stiffness, general health (GH) and worker productivity, and if so, the duration was determined. For unemployment and sick leave we used descriptive statistics.
Results: A flare negatively influenced GH, morning stiffness, HAQ-DI, EQ-5D, BRAF-MDQ, and the SF36 physical component scale and this effect lasted >3 months. Except for the HAQ-DI, effect sizes exceeded the minimum clinically important differences (MCIDs). For the physical outcomes effects lasted >6 months. Worker productivity was not significantly affected by a flare.
Conclusion: A disease flare influenced patients' lives, the largest effect was seen in the physical outcomes, and lasted 6 months. Although on a group level effect sizes for the separate PROMs were not always significant or larger than specific MCIDs, a disease flare can still be of great importance for individual patients.

Introduction

Over the years the treatment of Rheumatoid Arthritis (RA) has improved enormously, which resulted in better outcomes, including achievement of sustained remission [1,2]. Nowadays, 50–60% of RA patients achieve sustained remission [3,4]. Therefore, current guidelines recommend to consider tapering treatment if patients are in sustained remission [5,6].

Previous studies have shown that it is possible to taper biologicals, but this is accompanied with a higher chance of disease flares [7-10]. Flare rates within these studies varied from 38% to 76.6%. It has also been shown that only 41–67% of the patients that experienced a flare will regain remission within 6 months after treatment intensification [7,11,12]. Thus, many patients will have a reduced or no response to previous effective therapy, which may lead to an altered disease state or prolonged flare duration. Despite the high flare rates, current guidelines recommend to taper biologicals, which is based on a clinical and societal viewpoint.

At present, a paradigm shift in the delivery of health care is emerging, and is shifting towards patient centered healthcare. Patient-centered healthcare focuses on the individual patient preferences and needs, which can be objectified with patient reported outcome measures (PROMs) [13,14]. In order to optimize the delivery of care during tapering we need to know how a disease flare affects these PROMs. However, data on the feasibility of tapering DMARDs from a patient's perspective are sparse.

Abbreviations: ACPA, anti-citrullinated protein antibody; BRAF-MDQ, Bristol Rheumatoid Arthritis Fatigue multidimensional questionnaire; CRP, C-reactive protein; csDMARDs, conventional synthetic DMARD; DAS, disease activity score; ESR, erythrocyte sedimentation rate; EQ5D, European Quality of Life – 5 Dimensions; HADS, hospital anxiety and depression scale; HAQ-DI, health assessment questionnaire disability index; MCID, minimal clinically important difference; MCS, mental component scale; PCS, physical component scale; PROM, patient reported outcome measure; RF, rheumatoid factor; SF36, short form 36; SJC, swollen joint count; TJC, tender joint count; VAS, visual analogue scale

* Corresponding author.
E-mail address: elise.vanmulligen@erasmusmc.nl (E. van Mulligen).

https://doi.org/10.1016/j.semarthrit.2020.02.011

Therefore, our objectives are (1) to determine the impact of a disease flare on patient's lives by quantifying the changes in functional ability, general health, morning stiffness, fatigue, quality of life, and worker productivity, and (2) to explore the duration of this effect.

Methods

Study design

Data were used from the Tapering strategies in Rheumatoid Arthritis (TARA) trial (NTR2754). Adult patients with well-controlled RA, defined as a disease activity score (DAS44)≤2.4 and a swollen joint count (SJC)≤1 for at least 6 months, who were using a combination of a conventional synthetic disease modifying anti-rheumatic drug (csDMARD) and a TNF-inhibitor, were included. Patients were randomized into gradually tapering the csDMARD or TNF-inhibitor first. In the second year, the other drug was gradually tapered. The protocol was terminated if patients experienced a flare (DAS>2.4 and/or SJC>1). The previous effective dose was restarted and if necessary, medication was intensified further according to a treat-to-target approach, until low disease activity was reached. After a disease flare it was not allowed to restart tapering [12].

For the current study we compared the PROMs and DAS44, within all patients that experienced a flare, at the moment of flare, 3 months prior to a flare, and every 3 months thereafter with their own norm values. The norm was set at the average of DAS44 and PROMs 12, 9 and 6 months prior to a flare, which in our opinion was the best reference for well-controlled disease (Fig. 1).

We also performed a sensitivity analysis with different flare criteria from other studies, which are less strict than our criteria, in order to assess the impact of different criteria on measured outcomes. For example, we could have classified someone as having a disease flare, while in other studies these patients would continue tapering.

Outcomes

Outcomes for the impact of a disease flare on patients' lives were DAS, general health (GH), severity of morning stiffness, functional ability, quality of life, health status, fatigue, anxiety and depression, and worker productivity.

Every three months the DAS44 and self-reported questionnaires were collected [12]. The DAS44 was used for measuring disease activity based on 44 joints [15]. The minimum clinically important difference (MCID) of the DAS44 is 0.6 [16]. GH was measured on a 0–100 mm visual analogue scale, in which 0 represented the lowest possible health state, and 100 perfect health. The MCID for GH is 10 [16]. Functional ability was measured with the health assessment questionnaire disability index (HAQ-DI) [17]. Higher scores reflect greater disability, and the MCID is 0.22 [18]. Severity of morning stiffness was measured on a 0–10 likert-scale, in which 0 represented no morning stiffness, and 10 severe morning stiffness. The MCID for morning stiffness is 1 [16]. Quality of life (QoL) was measured with the European Quality of Life – 5 Dimensions (EQ-5D). Higher scores indicate a higher quality of life, and the MCID is 0.04 [19]. Health status was measured with the short form 36 (SF36), the higher the score, the better the health status [20-22]. The MCID of the SF36 is between 3 and 5 [22]. Fatigue was measured with the Bristol Rheumatoid Arthritis Fatigue Multi-dimensional Questionnaire (BRAF-MDQ). Higher scores represent higher levels of fatigue [23]. The MCID is 2.6 [24]. Anxiety and depression were measured with the Hospital Anxiety and Depression Scale (HADS), in which higher scores represent more anxiety and/or depression [25]. The HADS MCID for RA patients is unknown, however other chronic diseases show an MCID of 1.7 [26,27]. Worker productivity was assessed with the iMTA Productivity Cost Questionnaire (iPCQ) that addresses sick leave, reduction in work time, and productivity loss [28]. For all outcomes, the effect sizes were compared to aforementioned MCIDs.

Statistical analysis

We used data from patients that experienced a flare to determine the impact and duration of a flare on DAS44 and PROMs. The moment of flare was set as T0 and we only took the first flare into account. We used Linear Mixed Models (LMMs) with a random intercept and an autoregressive covariance matrix, to account for repeated measurements within individuals, to compare DAS44 and PROMs 3 months prior to a flare, at the moment of flare, and 3, 6, 9, and 12 months after a flare with norm values. For each patient the norm was set at the average value of DAS44 and PROMs for the combined values obtained at 12, 9 and 6 months prior to a flare. This was based on the mean DAS44 graph that showed minimal fluctuations between aforementioned timepoints in patients who experienced a flare and at those time-points these patients still had a well-controlled disease (Fig. 1). Because of aforementioned reasoning we had to exclude 17 patients, because they experienced a flare within the first 3 months of follow-up and, therefore, we could not set a norm value for these patients.

First, we examined whether there was a difference in each PROM and DAS44 over time. If there was a significant difference, the duration of this effect was determined. The duration was calculated by comparing each time-point separately with the norm, using aforementioned LMMs. For worker productivity we used descriptive statistics.

For visualization purposes, we also plotted the patients that did not have a disease flare. In this group we reclassified the 12 month visit as the new T0, because mean (sd) time to flare was 12 (6.7) months.

Outcomes were calculated in an intention-to-treat analysis, using all available data. A Bonferroni correction was applied to account for multiple testing. The calculated p-values for the impact of a flare on PROMs or DAS44 were corrected by multiplying the p-value with the total number of variables tested (n=11). The calculated p-values for the duration of a disease flare were multiplied with the total number of measurements tested (n=42). In this manner we could still consider a p-value ≤0.05 statistically significant. Corrected and uncorrected p-values are reported. All data were analyzed using STATA 15.

Results

Patients

A total of 189 patients were randomized, of those 113 patients experienced a flare. Table 1 shows the norm values for patients with and without a flare. Disease characteristics and PROMs were the same for both groups, except for DAS44 (sd), which was 0.86 (0.50) in the non-flare group and 1.08 (0.52) in the flare group (p=0.0055). This difference is probably caused by a significant difference in Erythrocyte Sedimentation Rate (ESR) between both groups (p=0.008).

Clinical outcomes

At the moment of flare (DAS44>2.4 or SJC>1), mean (sd) DAS44 was higher in the flare group (1.84 [0.76]) compared to the non-flare group (1.04 [0.51]) (Fig. 1A). Most of the separate components of the DAS44; TJC44, SJC44, and general health GH); were also higher in the flare group (Fig. 1B, C, E). We found an overall significant effect for the DAS44 compared to the norm (p<0.0001, Table 2). The same accounted for the DAS44 components, namely GH (p<0.0001, Table 2), SJC44 (p<0.0001), TJC44 (p<0.0001), ESR (p<0.0001), and CRP (p<0.0007) (data not shown). The effect of a flare on DAS44 and

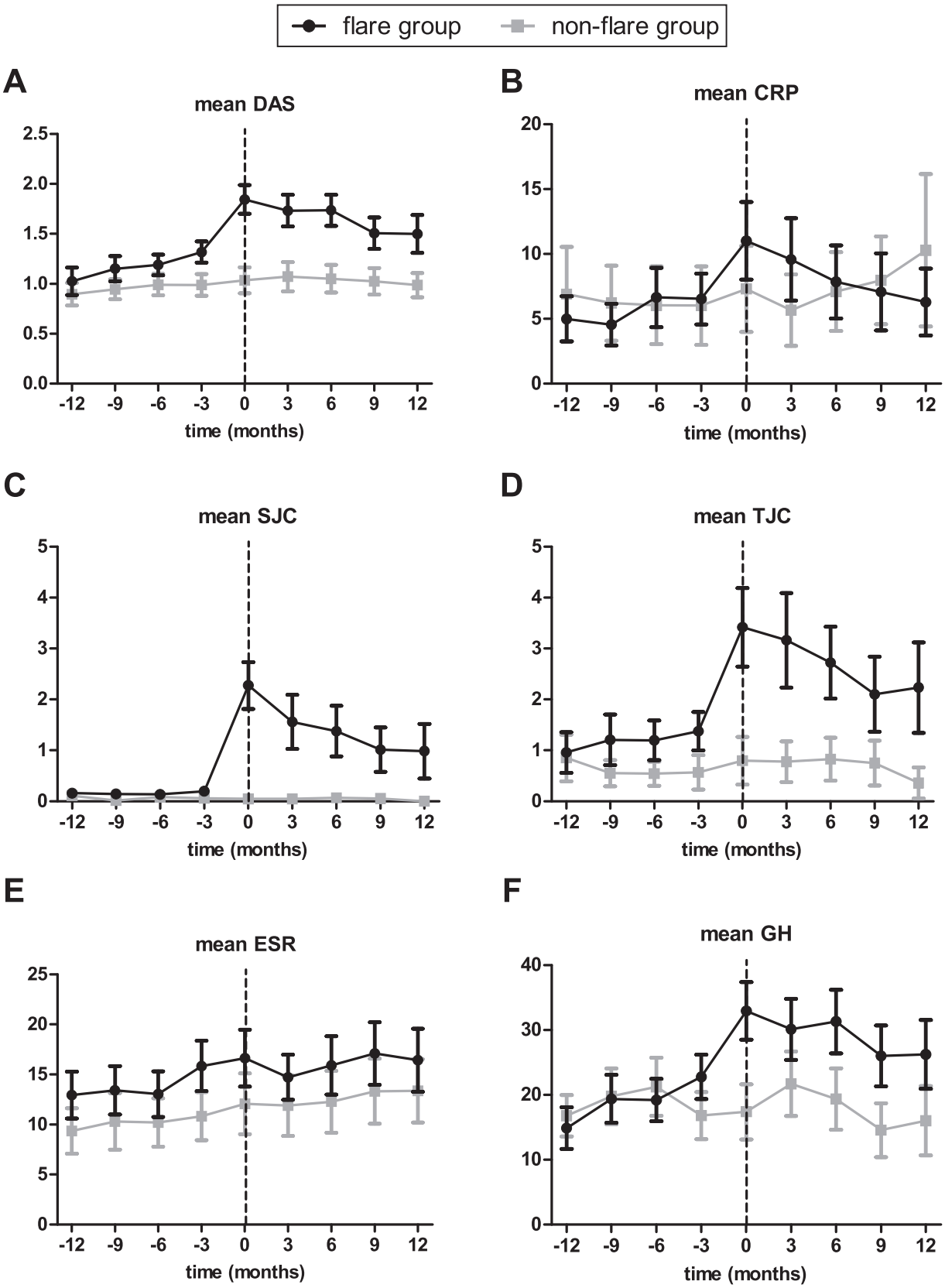

Fig. 1. Clinical outcomes. (A) DAS44 scores for the flare group and the non-flare group with corrected time-points. (B) Mean CRP, (C–F) separate components of the DAS44 scores: mean swollen joint count in 44 joints (SJC44), mean tender joint count in 44 joints (TJC44), mean erythrocyte sedimentation rate (ESR), and visual analogue scale for general health (GH).

Table 1
Patient characteristics.

Characteristics	Patients with flare (n=113)	Patients without flare (n=76)	p-value
Demographic at moment of randomization			
▪ Age (years), mean (sd)	58.2 (12.0)	54.1 (12.8)	0.025
▪ Gender, female, n (%)	77 (68.1)	48 (63.2)	0.48
Disease characteristics at moment of randomization			
▪ Symptom duration (years), median (IQR)	6.1 (4.3–9.1)	6.2 (3.8–8.5)	0.42
▪ RF positive, n (%)	61 (58.7)	45 (63.4)	0.53
▪ Erosive disease on initial radiograph, n (%)[a]	40 (42)	29 (38)	0.64
▪ ACPA positive, n (%)	75 (72.8)	52 (74.3)	0.83
Treatment at moment of randomization			
▪ MTX, n (%)	106 (94)	68 (89)	0.28
▪ Anti-TNF, n (%)			
- Etanercept	65 (58)	38 (50)	0.31
- Adalimumab	43 (38)	33 (43)	0.46
Norm values			
Disease activity			
▪ DAS44, mean (sd)	1.08 (0.52)	0.86 (0.50)	0.0055
▪ TJC44, median (IQR)	0 (0–1)	0 (0–0)	0.16
▪ SJC44, median (IQR)	0 (0–0)	0 (0–0)	0.12
▪ General health (0–100 mm), median (IQR)	14 (5–27)	14 (2–25.5)	0.82
▪ ESR (mm/h), median (IQR)	9.5 (5–16)	6 (2–12)	0.008
▪ CRP (mg/L), median (IQR)	2 (1–5)	2 (1–5.2)	0.73
▪ Morning stiffness, severity 0–10, median (IQR)	1 (0–3)	1 (0–4)	0.46
Patient reported outcomes			
▪ HAQ-DI, median (IQR)	0.38 (0.13–0.75)	0.25 (0–0.63)	0.57
▪ EQ-5D index, mean (sd)	0.86 (0.12)	0.87 (0.12)	0.51
▪ BRAF-MDQ, mean (sd)	16.2 (11)	16.6 (12)	0.80
▪ SF36, mean (sd)			
○ PCS	42.1 (11)	41.6 (11)	0.79
○ MCS	56.6 (10)	56.4 (9.0)	0.91
▪ HADS, mean (sd)	3.6 (3.0)	4.0 (2.8)	
○ Anxiety	2.0 (1.9)	2.7 (3.0)	0.39
○ Depression			0.08
▪ Worker productivity (0–10), median (IQR)	8 (6–10)	8 (6–9)	0.14

[a] Erosive disease is characterized as having >1 erosion in three separate joints. ACPA: anti-citrullinated protein antibody; CRP: C-reactive protein; csDMARDs: conventional synthetic DMARD; DAS: disease activity score; ESR: erythrocyte sedimentation rate; EQ5D: European Quality of Life – 5 Dimensions; HAQ-DI: health assessment questionnaire; IQR: inter quartile range; MCS: mental component scale; PCS: physical component scale; RF: rheumatoid factor; sd: Standard Deviation; SJC: swollen joint count; TJC: tender joint count.

GH lasted >12 months, while the clinically meaningful effect lasted 6 months (MCID DAS44>0.6 and MCID GH>10) [16].

The degree of morning stiffness, ranging from 0 to 10, was on average 3.7 (sd 2.8) in the flare group, and 2.5 (sd 2.3) in the non-flare group at T0 (Fig. 2B). The degree of morning stiffness significantly differed over time (p<0.0001, Table 2). When comparing the separate time-points to the norm, we found that morning stiffness significantly worsens at the moment of flare and regains its norm value 9 months after a flare. At the moment of flare and 3 months thereafter the difference with the norm was also above the MCID of 1 (Table 3) [16].

Functional ability

Functional ability was 0.69 (sd 0.61) at T0 in the flare group, and 0.47 (sd 0.56) in the non-flare group (Fig. 2B). When we visually compare the flare and non-flare group, we observed a difference that already starts six months prior to a flare and lasts until the end of the follow-up period. Not surprisingly the overall effect of a flare on the HAQ-DI was significant (p=0.0003, Table 2). However, when comparing the separate time-points to the norm, a significant difference was only observed at the moment of flare and 3 months thereafter. When taking uncorrected p-values into account, the effect would last longer, namely up to 9 months. However, the difference with norm values was never above the MCID of 0.22 (Table 3) [18].

Health status

For the health status we compared the flare group with the non-flare group based on the physical (PCS) and mental (MCS) component score of the SF36 (Fig. 2C, D). Mean PCS was 36.0 (sd 12.5) in the flare group and 42.8 (sd 11.0) in the non-flare group. The mean MCS was respectively 55.8 (sd 9.5) and 56.1 (sd 10.0) in the flare and non-flare group. The overall effect of flare was not significant for the MCS (p=1), but it was for the PCS (p=0.0004). If we compare the separate time-points to norm values, a significant effect was only present at the moment of flare (Table 3) [22]. Using the uncorrected p-values, there was a significant and also a clinically meaningful effect, which lasted up to 6 months after a flare (MCID SF36 PCS 3-5).

Quality of life

Quality of life shows a small dell in the graph at the moment of flare (Fig. 2E). The mean EQ-index at T0 was, respectively 0.75 (sd 0.21) and 0.85 (sd 0.13) for the patients who did and did not experienced a flare. The overall effect of a flare on EQ-5D was significant (p<0.0001, Table 2), which was also seen in the separate domains (p<0.01), except for the domain anxiety and depression (p=0.46, data not shown). This significant effect was only seen at the moment of flare, which also exceeded the MCID threshold of 0.04 [19]. If we look at the uncorrected p-values, there was a significant effect that lasted >12 months with an effect size ≥MCID for all significant time-points (Table 3) [19].

Fatigue

At T0 we encountered a mean fatigue score of 19.6 (sd 11.5) in the flare group and 15.7 (sd 13.1) in the non-flare group (Fig. 2F). The effect of a flare on fatigue was significant (p=0.042, Table 2). However, when comparing separate time-points the corrected p-values were not significant, while the uncorrected p-values showed a duration of 6 months. During this time period the difference with norm also exceeded the MCID of 2.6 (Table 3) [24].

Anxiety and depression

At visual inspection of the anxiety and depression graphs an erratic course of the scores is observed (Fig. 2G, H). At the moment of flare the mean anxiety scores were 3.63 (sd 2.89) and 3.25 (sd 2.96)

Table 2
Overall differences between norm and moments thereafter.

Patient reported outcomes	p-value	Bonferroni corrected p-value[a]
DAS44	p<0.0001	p<0.0001
VAS general health	p<0.0001	p<0.0001
Morning stiffness	p<0.0001	p<0.0001
HAQ-DI	p<0.0001	p=0.0003
SF36 PCS	p<0.0001	p=0.0004
SF36 MCS	p=0.68	p=1
EQ5D	p<0.0001	p<0.0001
BRAF-MDQ	p=0.0037	p=0.041
HADS anxiety	p=0.75	p=1
HADS depression	p=0.62	p=1
Worker productivity	p=0.32	p=1

[a] n=11. BRAF-MDQ: Bristol Rheumatoid Arthritis Fatigue multidimensional questionnaire; DAS: disease activity score; EQ5D: European Quality of Life with 5 Dimensions; HADS: hospital anxiety and depression scale; HAQ-DI: health assessment questionnaire disability index; MCS: mental component scale; PCS: physical component scale; SF36: short form 36; VAS: visual analogue scale.

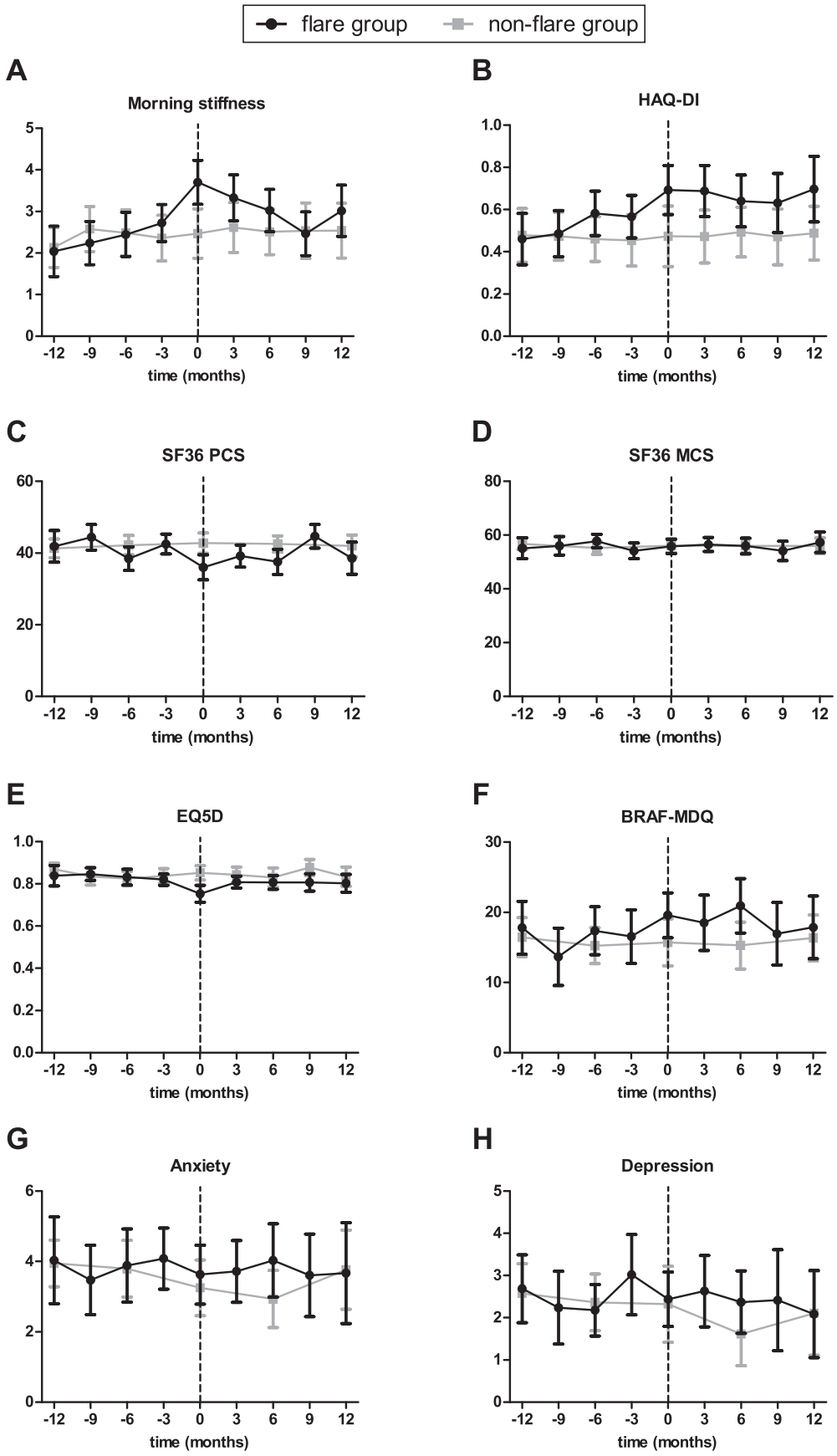

Fig. 2. Patient reported outcome measures (PROMs). EQ5D: European Quality of Life with 5 Dimensions; HAQ-DI: health assessment questionnaire disability index; SF36: short form 36; MCS: mental component scale; PCS: physical component scale.

Table 3
Comparison of separate time-points with the norm values to assess the duration of the effect of flare.

		Difference with norm (effect size)	95% CI			*p*-value	Bonferroni corrected *p*-value[a]
DAS44	-T3	0.16	0.039	–	0.27	0.0089	0.37
(MCID=0.6) [16]	T0	0.68	0.56	–	0.81	<0.0001	<0.0001
	T3	0.57	0.44	–	0.70	<0.0001	<0.0001
	T6	0.57	0.43	–	0.71	<0.0001	<0.0001
	T9	0.33	0.18	–	0.47	<0.0001	0.0004
	T12	0.32	0.16	–	0.47	0.0001	0.0027
General health	-T3	4.4	1.10	–	7.78	0.0091	0.38
(MCID=10) [16]	T0	14.8	11.18	–	18.35	<0.0001	<0.0001
	T3	12.4	8.53	–	16.21	<0.0001	<0.0001
	T6	12.7	8.65	–	16.74	<0.0001	<0.0001
	T9	7.6	3.34	–	11.92	0.0005	0.021
	T12	7.8	3.22	–	12.48	0.0009	0.037
Morning stiffness (MCID=1) [16]	-T3	0.41	0.028	–	0.78	0.036	1
	T0	1.32	0.93	–	1.72	<0.0001	<0.0001
	T3	1.15	0.73	–	1.57	<0.0001	<0.0001
	T6	0.86	0.41	–	1.30	0.0001	0.0062
	T9	0.26	−0.21	–	0.73	0.28	1
	T12	0.87	0.36	–	1.37	0.0007	0.031
HAQ-DI	-T3	0.016	−0.039	–	0.071	0.57	1
(MCID=0.22) [18]	T0	0.13	0.074	–	0 .19	<0.0001	0.0002
	T3	0.12	0.065	–	0.18	<0.0001	0.0019
	T6	0.078	0.015	–	0.14	0.015	0.61
	T9	0.046	−0.021	–	0.11	0.18	1
	T12	0.083	0.011	–	0.16	0.025	1
SF36 PCS	-T3	−1.03	−3.67	–	1.62	0.45	1
(MCID=3–5) [22]	T0	−4.25	−6.57	–	-1.93	0.0003	0.014
	T3	−4.05	−6.84	–	-1.27	0.0044	0.18
	T6	−3.95	−6.64	–	-1.26	0.0041	0.17
	T9	1.37	−1.71	–	4.45	0.38	1
	T12	−2.93	−6.11	–	0.26	0.072	1
EQ5D	-T3	−0.020	−0.048	–	0.0070	0.15	1
(MCID=0.04) [19]	T0	−0.086	−0.11	–	-0.059	<0.0001	<0.0001
	T3	−0.042	−0.071	–	-0.014	0.0039	0.16
	T6	−0.036	−0.067	–	-0.0064	0.018	0.73
	T9	−0.037	−0.069	–	-0.0052	0.023	0.95
	T12	−0.047	−0.081	–	-0.012	0.0081	0.34
BRAF-MDQ (MCID=2.6) [24]	-T3	1.41	−1.22	–	4.03	0.29	1
	T0	3.15	0.94	–	5.36	0.0053	0.68
	T3	3.25	0.60	–	5.90	0.016	1
	T6	4.33	1.85	–	6.82	0.0006	0.026
	T9	1.58	−1.32	–	4.49	0.29	1
	T12	1.76	−1.14	–	4.66	0.23	1

[a] *n*=42. BRAF-MDQ: Bristol Rheumatoid Arthritis Fatigue multidimensional questionnaire; CI: confidence interval; DAS: disease activity score; EQ5D: European Quality of Life with 5 Dimensions; HADS: hospital anxiety and depression scale; HAQ-DI: health assessment questionnaire disability index; MCID: minimal clinically important difference; MCS: mental component scale; PCS: physical component scale; SF36: short form 36; VAS: visual analogue scale.

for the flare and non-flare group. Mean depression scores were respectively 2.44 (sd 2.22) and 2.32 (sd 3.36) for the flare and non-flare group. Depression as well as anxiety scores were not influenced by a flare (*p*=1 for both scores, Table 2).

Worker productivity

We first determined how many patients had payed work (Fig. 3A). At T0, 48% of the flare group and 59% of the non-flare group had payed work. Over time there were only minor differences in these numbers. Of the eligible working population respectively 27% and 18% of patients with and without a flare were unemployed at T0 (Fig. 3B). These percentages did not vary much over time. Sick leave was 6.2% in the flare group, and 2.6% in the non-flare group at T0, which was measured over the entire working population (Fig. 3C). Sick leave was not clearly affected by a flare, although we did saw a 10% drop in productivity in the 3 months after a flare, which was not significant (Fig. 3).

Sensitivity analysis

We performed a sensitivity analysis to evaluate the effect of different flare criteria on our PROMs (Supplementary Table S1). For all flare definitions we found that DAS44, GH, morning stiffness, HAQ-DI and the EQ5D were affected (Supplementary Table S1). The effect of these different flare definitions on PROMs might even be larger compared to our results.

Discussion

We showed that a disease flare has a significant effect on all components of the disease activity score, but also on functional ability, quality of life and fatigue, which lasted at least 3 months. Worker productivity did not seem to be affected by a flare.

In the TARA study it was shown that tapering csDMARDs or anti-TNF in established RA patients resulted in an average flare rate of 38% during the first year of follow-up. The two tapering arms did not differ in flare rates, functional ability or quality of life [29]. Six months after the flare, 67% of the patients regained well-controlled disease [29]. These results were comparable with other tapering studies [7,9,11,30]. In all these studies PROMs were merely not taken into account to assess the severity of a flare. Furthermore, it was not investigated if PROMs differed between patients with and without a flare.

However, the POET trial did show that stopping the TNF-inhibitor had a significant short-term impact on physical and mental health

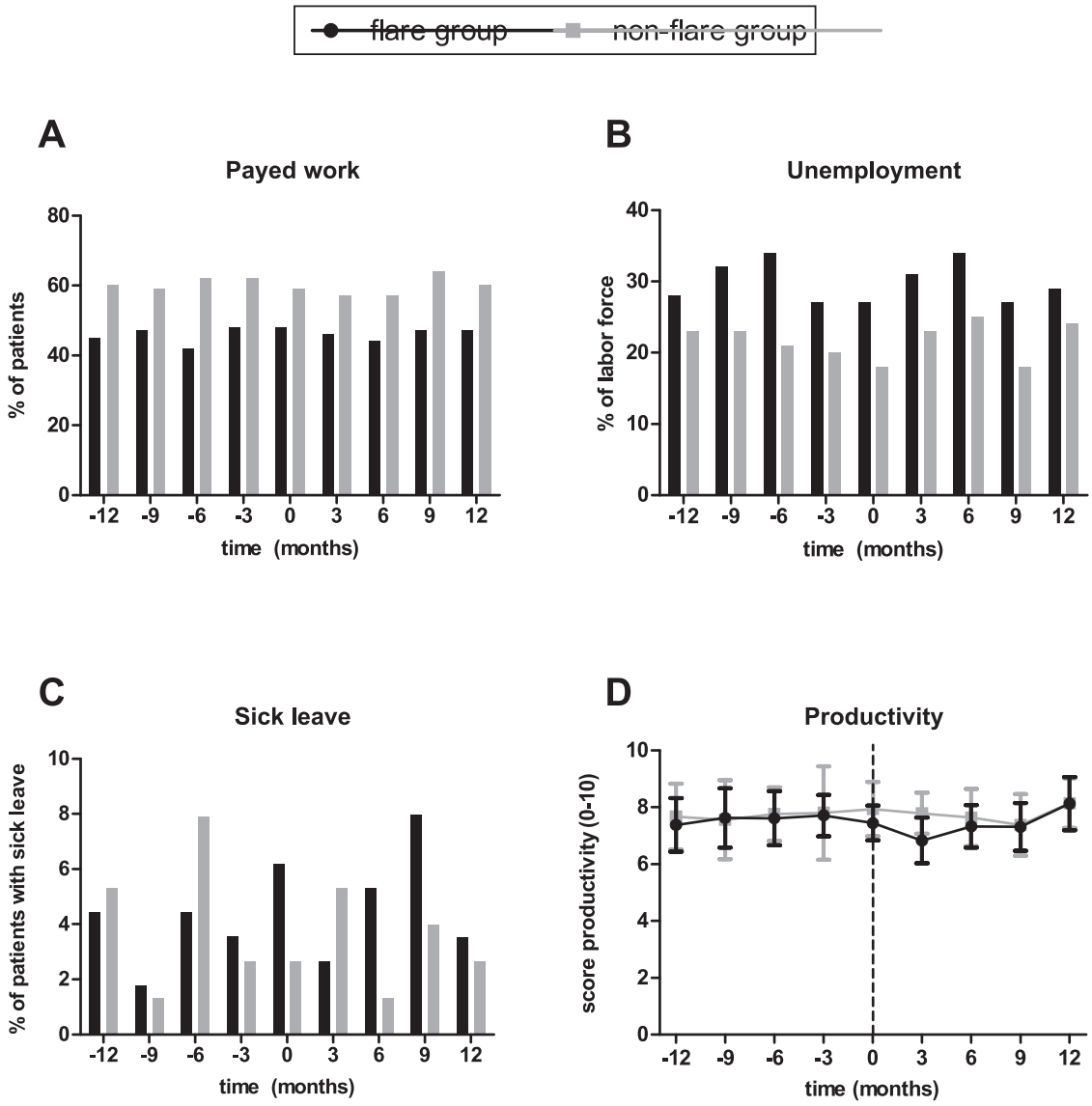

Fig. 3. Worker productivity. (A) The percentage of patients with payed work, (B) unemployment as a percentage of the total labor force, (C) the amount of sick leave indicated as number of patients calling in sick within a 3 month period, (D) productivity on a scale from 0 to 10.

status compared to patients who continued their TNF-inhibitor [31]. Furthermore, the STRASS trial investigated whether the patient's perspective of a flare was the same as the physician's perspective of a flare, which was measured with the DAS28 [32]. The investigators concluded that the patient reported flare overlapped with the DAS28-based flare. The OPTIRRA trial investigators explored whether PROMs could predict a flare [33]. They showed that mental health status was independently associated with a flare during tapering. Also fatigue and functional ability were associated with a flare, but this effect disappeared after correction for possible confounders.

Although we showed a significant effect of a flare on various PROMs, this effect was not always above the MCID. For the HAQ-DI, for example, the MCID is 0.22, which was not reached in our analysis. However, the differences with the norm were statistically significant up to 6 months after a flare. Not reaching the MCID, while finding a significant differences, might be due to our assumption for the norm values. The norm was set at the average of the visits 12, 9 and 6 months prior to a flare, which was based upon the DAS44 graph. If we look at the HAQ-DI graph, we see that the HAQ-DI already worsens 6 months prior to a flare. Therefore, by taking this visit as part of the norm value, we might have underestimated the effect of flare on the HAQ-DI. For the EQ5D and the SF36 PCS we can apply a reverse reasoning of the foregoing explanation. For both PROMs we only found a significant difference at the moment of flare, while the MCID was reached for almost every time-point after the flare, which indicates that a disease flare might have great impact on individual patients. Moreover, we corrected for multiple testing, which might have canceled out a possible meaningful effect and, therefore, underestimated the significance of our results. On the other hand, our sensitivity analysis showed similar findings for different flare definitions, which strengthens our current findings.

Strengths of the current study include the completeness of the data, including containment of recommended outcomes measures by ICHOM and OMERACT [14,34]. Furthermore the TARA trial used a gradual tapering scheme combined with a treat-to-target approach. Therefore, we think this is an ideal trial to investigate the effect of a flare on PROMs.

Limitations of this study were that it is a post-hoc analysis. However, due to our statistical approach in which we compared patients with their own norm values, we think we can still report valid results. The results on worker productivity on the other hand are less reliable, because of the low occurrence of absenteeism and presenteeism,

giving rise to a potential power issue. Furthermore, the TARA trial only had a follow-up period of 2 years, whereby potential long term effects could not be determined. For some of the investigated PROMs we already saw a long lasting effect (>6 months). Ideally, we would like to know exactly how long aforementioned effects are present, but unfortunately we do not have the data for this. There is also not always consensus about the MCIDs for specific PROMs. We used known MCIDs from the literature to place our result into perspective, but it is debatable if those values are correct.

Recently, there has been some debate on the measurement of morning stiffness, and efforts are made to create a validated PROM according to OMERACT guidelines [35,36]. Current used measures do not capture all aspects that are involved with morning stiffness due to RA disease activity. However, the OMERACT working group does advice not to use morning stiffness duration as outcome, because it is very aspecific [35]. Fortunately, we used the severity of morning stiffness as outcome in our analyses, but one should be cautious when interpreting these outcomes.

Due to the long-lasting effect of a flare on a patient's live, it would be ideal if we were able to predict who can safely taper medication. Current tapering strategies are based upon a trial-and-error approach, which leads to high flare rates. Our study showed that some PROMs already worsen before a flare occurs, i.e. HAQ-DI, severity of morning stiffness and the DAS44, which might be useful for flare prevention during tapering. These changes before the actual flare occurred were all non-significant, still it indicates that patients already have more complaints before the actual flare was objectified by the treating physician. Therefore, the results of this study could be used for future research to establish a more personalized tapering approach, even though prediction of flares is not yet possible.

In conclusion, a disease flare has a significant effect on patients' lives. A disease flare affects functional ability, quality of life, fatigue, and all components of the disease activity score. The largest effect was seen in the physical outcomes, and lasted 6 months. Although on a group level the effect size for several PROMs did not exceed the specific MCID, a disease flare can still be of great importance for individual patients.

Declaration of Competing Interest

None.

Funding

This work was supported by ZonMw [grant number 171102014] and Pfizer [grant number WP1696714].

Acknowledgments

We especially thank the participating patients in the TARA trial for their willingness to contribute to the study and for their cooperation. We thank all rheumatologists from the following participating centers: Erasmus MC, Rotterdam; Maasstad ziekenhuis, Rotterdam; Sint Fransiscus Gasthuis & Vlietland, Rotterdam & Schiedam; Amphia ziekenhuis, Breda; Reinier de Graaf ziekenhuis, Delft; Antonius ziekenhuis, Utrecht; Bravis ziekenhuis, Bergen op Zoom & Roosendaal; Groene hart ziekenhuis, Gouda; Albert Schweitzer ziekenhuis, Dordrecht; Haga ziekenhuis, Den Haag; and Beatrix ziekenhuis, Gorinchem. We also thank all study-nurses, laboratory personnel, co-investigators and others who were involved with the TARA study and also all participating hospitals in which patients were treated and recruited.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.semarthrit.2020.02.011.

References

[1] Nieuwenhuis WP, de Wit MP, Boonen A, van der Helm-van Mil AH. Changes in the clinical presentation of patients with rheumatoid arthritis from the early 1990s to the years 2010: earlier identification but more severe patient reported outcomes. Ann Rheum Dis 2016;75(11):2054–6.
[2] Haugeberg G, Boyesen P, Helgetveit K, Proven A. Clinical and radiographic outcomes in patients diagnosed with early rheumatoid arthritis in the first years of the biologic treatment era: a 10-year prospective observational study. J Rheumatol 2015;42(12):2279–87.
[3] Lenert A, Lenert P. Tapering biologics in rheumatoid arthritis: a pragmatic approach for clinical practice. Clin Rheumatol 2017;36(1):1–8.
[4] Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ. Flare rate in patients with rheumatoid arthritis in low disease activity or remission when tapering or stopping synthetic or biologic DMARD: a systematic review. J Rheumatol 2015;42(11):2012–22.
[5] Singh JA, Saag KG, Bridges Jr. SL, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016;68(1):1–26.
[6] Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, 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(6):960–77.
[7] Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, et al. Stopping tumor necrosis factor inhibitor treatment in patients with established rheumatoid arthritis in remission or with stable low disease activity: a pragmatic multicenter, open-label randomized controlled trial. Arthritis Rheumatol 2016;68(8):1810–7.
[8] Fautrel B, den Broeder AA. De-intensifying treatment in established rheumatoid arthritis (RA): why, how, when and in whom can DMARDs be tapered? Best Pract Res Clin Rheumatol 2015;29(4):550–65.
[9] van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, et al. Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. Cochrane Database Syst Rev 2014(9):CD010455.
[10] Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, et al. Relapse rates in patients with rheumatoid arthritis in stable remission tapering or stopping antirheumatic therapy: interim results from the prospective randomised controlled RETRO study. Ann Rheum Dis 2016;75(1):45–51.
[11] Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, et al. Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). Ann Rheum Dis 2016;75(1):59–67.
[12] van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, et al. Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. Ann Rheum Dis 2019;78(6):746–53.
[13] Heller JE, Shadick NA. Outcomes in rheumatoid arthritis: incorporating the patient perspective. Curr Opin Rheumatol 2007;19(2):101–5.
[14] Oude Voshaar MAH, Das Gupta Z, Bijlsma JWJ, Boonen A, Chau J, Courvoisier DS, et al. The International Consortium for Health Outcome Measurement (ICHOM) set of outcomes that matter to people living with inflammatory arthritis consensus from an international working group. Arthritis Care Res 2019;71(12):1556–65.
[15] van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49(11):916–20.
[16] Anderson JK, Zimmerman L, Caplan L, Michaud K. Measures of rheumatoid arthritis disease activity: Patient (PtGA) and Provider (PrGA) Global Assessment of Disease Activity, Disease Activity Score (DAS) and Disease Activity Score with 28-Joint Counts (DAS28), Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), Patient Activity Score (PAS) and Patient Activity Score-II (PASII), Routine Assessment of Patient Index Data (RAPID), Rheumatoid Arthritis Disease Activity Index (RADAI) and Rheumatoid Arthritis Disease Activity Index-5 (RADAI-5), Chronic Arthritis Systemic Index (CASI), Patient-Based Disease Activity Score With ESR (PDAS1) and Patient-Based Disease Activity Score without ESR (PDAS2), and Mean Overall Index for Rheumatoid Arthritis (MOI-RA). Arthritis Care Res 2011;63(Suppl 11):S14–36.
[17] Siegert CE, Vleming LJ, Vandenbroucke JP, Cats A. Measurement of disability in Dutch rheumatoid arthritis patients. Clin Rheumatol 1984;3(3):305–9.
[18] Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Arch Intern Med 1993;153(11):1337–42.
[19] Luo N, Johnson J, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. Med Care 2010;48(4):365–71.
[20] EuroQol G. EuroQol – a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199–208.
[21] Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997;36(5):551–9.
[22] Ware Jr. JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30(6):473–83.
[23] Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Collaboration with patients in the design of patient-reported outcome measures: capturing the experience of fatigue in rheumatoid arthritis. Arthritis Care Res 2010;62(11):1552–8.

[24] Dures EK, Hewlett SE, Cramp FA, Greenwood R, Nicklin JK, Urban M, et al. Reliability and sensitivity to change of the bristol rheumatoid arthritis fatigue scales. Rheumatology 2013;52(10):1832–9.
[25] Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67(6):361–70.
[26] Puhan MA, Frey M, Buchi S, Schunemann HJ. The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. Health Qual Life Outcomes 2008;6:46.
[27] Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the minimal clinically important difference for the hospital anxiety and depression scale in patients with cardiovascular disease. J Cardiopulm Rehabil Prev 2019;39(6):E6–E11.
[28] Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. The iMTA productivity cost questionnaire: a standardized instrument for measuring and valuing health-related productivity losses. Value Health 2015;18(6):753–8.
[29] van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, et al. Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. Ann Rheumatic Dis 2019;78(6):746–53.
[30] Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, et al. Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. Lancet 2013;381(9870):918–29.
[31] Ghiti Moghadam M, Ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, et al. Impact of stopping tumor necrosis factor inhibitors on rheumatoid arthritis patients' burden of disease. Arthritis Care Res 2018;70(4):516–24.
[32] Portier A, Gossec L, Tubach F, Alfaiate T, Pham T, Saraux A, et al. Patient-perceived flares in rheumatoid arthritis: a sub-analysis of the STRASS treatment tapering strategy trial. Joint Bone Spine 2017;84(5):577–81.
[33] Bechman K, Sin FE, Ibrahim F, Norton S, Matcham F, Scott DL, et al. Mental health, fatigue and function are associated with increased risk of disease flare following TNF inhibitor tapering in patients with rheumatoid arthritis: an exploratory analysis of data from the Optimizing TNF Tapering in RA (OPTTIRA) trial. RMD Open 2018;4(1):e000676.
[34] van Tuyl LH, Sadlonova M, Davis B, Flurey C, Goel N, Hewlett SE, et al. Remission in rheumatoid arthritis: working toward incorporation of the patient perspective at OMERACT 12. J Rheumatol 2016;43(1):203–7.
[35] Craig ET, Orbai AM, Mackie S, Bartlett SJ, Bingham 3rd CO, Goodman S, et al. Advancing stiffness measurement in rheumatic disease: report from the stiffness special interest group at OMERACT 2018. J Rheumatol 2019;46(10):1374–8.
[36] Halls S, Sinnathurai P, Hewlett S, Mackie SL, March L, Bartlett SJ, et al. Stiffness is the cardinal symptom of inflammatory musculoskeletal diseases, yet still variably measured: report from the OMERACT 2016 stiffness special interest group. J Rheumatol 2017;44(12):1904–10.