Arthritis Care & Research
Vol. 70, No. 4, April 2018, pp 516–524
DOI 10.1002/acr.23315


ORIGINAL ARTICLE

# Impact of Stopping Tumor Necrosis Factor Inhibitors on Rheumatoid Arthritis Patients' Burden of Disease

MARJAN GHITI MOGHADAM,[1] PETER M. TEN KLOOSTER,[1] HARALD E. VONKEMAN,[1] EVA L. KNEEPKENS,[2] RUTH KLAASEN,[3] JAN N. STOLK,[4] ILJA TCHETVERIKOV,[5] SIMONE A. VREUGDENHIL,[6] JAN M. VAN WOERKOM,[7] YVONNE P. M. GOEKOOP-RUITERMAN,[8] ROBERT B. M. LANDEWÉ,[9] PIET L. C. M. VAN RIEL,[10] MART A. F. J. VAN DE LAAR,[1] AND TIM L. JANSEN,[11] ON BEHALF OF THE DUTCH NATIONAL POET COLLABORATION

***Objective***. **To determine the impact of stopping tumor necrosis factor inhibitor (TNFi) treatment on patient-reported outcomes (PROs) of physical and mental health status, health utility, pain, disability, and fatigue in patients with established rheumatoid arthritis (RA).**

***Methods***. **In the pragmatic, 12-month POET trial, 817 RA patients with ≥6 months of remission or stable low disease activity were randomized 2:1 to stopping or continuing TNFi. In case of flare, TNFi was restarted at the discretion of the rheumatologist. PROs were assessed every 3 months.**

***Results***. **TNFi was restarted within 12 months in 252 of 531 patients (47.5%) in the stop group. At 3 months, mean PRO scores were significantly worse in the stop group, and a larger proportion of patients experienced a minimum clinically important difference (MCID) on all PROs. Effect sizes (ES) were strongest for health utility (ES −0.24) and pain (ES −0.30). Mean scores improved again after this point, but disability scores remained significantly different at 12 months. After 12 months, the relative risk of experiencing an MCID ranged from 1.16 for mental health status to 1.58 for fatigue. Mean PRO scores for patients restarting TNFi within 6 months were no longer significantly different from those that did not restart TNFi at 12 months.**

***Conclusion***. **Stopping TNFi had a significant negative short-term impact on a broad range of PROs. Long-term negative consequences appeared to be limited, and outcomes in patients needing to restart TNFi within the first 6 months tended to be restored at 12 months.**

## INTRODUCTION

The introduction of biologic agents such as tumor necrosis factor inhibitors (TNFi) has drastically improved the outcomes and clinical course of rheumatoid arthritis (RA). Their efficacy and effectiveness has been extensively demonstrated (1–3), and several trials have additionally demonstrated rapid and sustained improvements in a range of patient-reported outcomes (PROs), including health-related quality of life, general health, pain, disability, and fatigue (4–8).

As the long-term use of TNFi is costly and associated with increased risk of side effects (9–11), the possibility of stopping TNFi in patients with stable low disease activity (LDA) or remission is receiving increasing interest. Several retrospective studies have suggested that stopping TNFi is

Netherlands Trial Register: NTR3112.

Supported by The Netherlands Organization for Health Research and Development (ZonMw)/Government of The Netherlands, Ministry of Health, Welfare and Sport.

[1]Marjan Ghiti Moghadam, MD, Peter M. ten Klooster, PhD, Harald E. Vonkeman, MD, PhD, Mart A. F. J. van de Laar, MD, PhD: Arthritis Centre Twente, University of Twente, and Medisch Spectrum Twente, Enschede, The Netherlands; [2]Eva L. Kneepkens, MD, PhD: Reade, Amsterdam, The Netherlands; [3]Ruth Klaasen, MD, PhD: Meander Medical Centre, Amersfoort, The Netherlands; [4]Jan N. Stolk, MD, PhD: Gelderse Vallei Hospital, Ede, The Netherlands; [5]Ilja Tchetverikov, MD, PhD: Albert Schweitzer Hospital, Dordrecht, The Netherlands; [6]Simone A. Vreugdenhil, MD: St. Antonius Hospital, Nieuwegein, The Netherlands; [7]Jan M. van Woerkom, MD, PhD: Gelre Hospitals, Apeldoorn, The Netherlands; [8]Yvonne P. M. Goekoop-Ruiterman, MD, PhD: Haga Hospital, The Hague, The Netherlands; [9]Robert B. M. Landewé, MD, PhD: Academic Medical Center, Amsterdam, The Netherlands; [10]Piet L. C. M. van Riel, MD, PhD: Radboud University Medical Center, Nijmegen, The Netherlands; [11]Tim L. Jansen, MD, PhD: VieCuri Medical Center, Venlo, The Netherlands.

Address correspondence to Marjan Ghiti Moghadam, MD, University Hospital Leuven, Department of Rheumatology, Herestraat 49, 3000 Leuven, Belgium. E-mail: m.ghiti@hotmail.com.

Submitted for publication February 7, 2017; accepted in revised form July 6, 2017.

## Significance & Innovations

- This is the first large study to determine the effects of stopping tumor necrosis factor inhibitor (TNFi) treatment on a broad range of patient-reported outcomes (PROs).
- Stopping TNFi resulted in a substantial short-term worsening on all outcomes, but mean 12-month outcome scores were comparable with those in the continuation group.
- Mean PROs in patients restarting TNFi after a flare in the first 6 months tended to normalize at 12 months.

possible in some patients (12,13), although 2 randomized trials showed that the majority of patients experienced a disease relapse within 1 year after stopping (14,15). This finding was recently confirmed in the POET trial (16), which showed that 51.2% of patients who stopped TNFi experienced a clinical flare within 12 months of followup compared to 18.2% of patients who continued TNFi treatment.

Current studies on stopping TNFi have primarily focused on the effects of discontinuation on disease activity status as the primary end point. However, clinical measures of disease activity may not adequately reflect the perception of illness and symptoms by RA patients themselves (17–20). Moreover, those studies that did report 1 or more PROs as secondary end points only examined group-level changes, providing no information on the consequences of stopping TNFi at the individual patient level (21). Consequently, little is currently known about the effects of trying to stop TNFi on the burden of disease as experienced by patients, especially in those patients who cannot successfully stop. The objectives of the current study were to determine the effects of stopping TNFi treatment on important PROs and to explore the course of symptoms in patients restarting TNFi after a flare.

## PATIENTS AND METHODS

**Patients.** Detailed study design and results of primary end points of the POET trial have been published (16). Briefly, the POET study was a pragmatic, open-label trial in which RA patients with stable LDA (defined as a Disease Activity Score in 28 joints [DAS28] of <3.2 for at least 6 months prior to inclusion) from 47 rheumatology centers throughout The Netherlands were randomized 2:1 to either stop or continue TNFi treatment. All participating patients were age ≥18 years, had RA per the American College of Rheumatology (ACR) 1987 criteria (22), and had received TNFi treatment for at least 1 year prior to inclusion. Concomitant treatment with conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) was continued. If RA flared, defined as a DAS28 score ≥3.2 with an increase >0.6 (23), TNFi could be restarted at the discretion of the treating rheumatologist. The study was approved by the ethical review boards of all participating hospitals and conducted in accordance with Good Clinical Practice guidelines and the Declaration of Helsinki. The POET study is registered in the Netherlands Trial Register (NTR3112).

In total, 531 patients were allocated to the stop group and 286 to the continuation group and followed for 12 months (16). The primary outcome of the study was that significantly more people in the stop group (272 of 531 [51.2%]) experienced a DAS28 flare than in the continuation group (52 of 286 [18.2%]). Most of these flares (213 of 272 [78.3%]) occurred within 6 months after stopping TNFi. TNFi was restarted within 6 months in 211 patients (39.7%) and within 12 months in 252 patients (47.5%) in the stop group. Almost 85% of the patients who restarted TNFi within the first 6 months had regained LDA at 12 months.

**Patient assessments.** Patients were evaluated by their treating rheumatologist and rheumatology nurse at baseline and at least once every 3 months thereafter, for a period of 1 year. Baseline measures included age, sex, weight, height, disease duration, medication use, rheumatoid factor (RF), and anti–citrullinated protein antibody status, as well as concomitant use of csDMARDs. Clinical measurements were performed at every scheduled or unscheduled visit and included the erythrocyte sedimentation rate (ESR, mm/hour), C-reactive protein level (mg/dl), tender joint count in 28 joints, swollen joint count in 28 joints, and a patient-reported assessment of general health on a 100-mm visual analog scale. These component measures were combined to calculate the DAS28-ESR (24). Physician-reported flares and all changes in medication were continuously recorded throughout the study.

Patients additionally completed 6 established PROs at baseline and before every study visit, including the 36-Item Short Form (SF-36) health survey (version 2.0) (25), the EuroQol 5-domain (EQ-5D) measure (26), the Health Assessment Questionnaire (HAQ) disability index (DI) (27), and the Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire (BRAF-MDQ) (28). Except for the BRAF-MDQ, well-established values for minimum clinically important differences (MCIDs) are available for each of these PROs as described below. Patients in the stop group also rated their overall experience of stopping TNFi before each study visit.

**Measures.** *Health status.* The SF-36 assesses different aspects of health represented in 8 scales (25,29). The 8 SF-36 scales are linearly transformed to range from 0 to 100, with higher scores representing better health status. Additionally, the scale scores can be aggregated into a physical component summary (PCS) and a mental component summary (MCS). The component summary scores are standardized using normative data from the 1998 US general population with a mean score of 50 and an SD of 10. Changes of 5 to 10 points on the 0–100 scales of the SF-36 and of 2.5 to 5 on the norm-based component scales are considered to be clinically meaningful (30,31). The data were analyzed based on an MCID of ≥10 points on the bodily pain (BP) scale and ≥5 points on the component summaries, as this corresponds closely with the half an SD rule of thumb (32).

*Health utility.* Health-related quality of life was additionally measured with the EQ-5D (33). The EQ-5D assesses problems in 5 domains (mobility, self-care, usual activities,

pain/discomfort, and anxiety/depression) on 3-point scales to produce a single, interval-level utility score. Utilities reflect the relative desirability of the health state, where 0 refers to death and 1 refers to full health. Utility scores were calculated using the Dutch tariff (34). The MCID for the EQ-5D in RA was reported to be 0.05 (35).

*Functional disability.* The HAQ DI contains 20 items measuring limitations over the past week in 8 categories of daily living (27,36). Each item is scored on a 4-point rating scale from 0 (without any difficulty) to 3 (unable to do). The standard total disability score was calculated by determining the highest score in each of the 8 categories, corrected for the use of aids and devices, and then averaging the category scores (37). Scores on the HAQ DI range between 0 and 3, with higher values indicating more disability. A 2-step difference in score, or 0.25 units, is considered clinically significant (38).

*Fatigue.* The BRAF-MDQ is a disease-specific, 20-item questionnaire covering domains of physical fatigue (4 items), living (7 items), cognition (5 items), and emotion (4 items) (28). All 20 items can be summed to produce a global fatigue score (range 0–70), with higher scores representing worse fatigue. As no formal MCID has been established for the BRAF-MDQ, a minimal increase of 7 points was considered to indicate clinically relevant worsening, which corresponds to approximately half an SD as reported in previous studies (28,39,40).

*Patient-reported experience of stopping TNFi.* Patients in the stop group rated their experience of stopping TNFi on a single item ("I experience stopping with the TNF blocker as…") with a Likert-type response scale ranging from 1 (very positive) to 5 (very negative).

**Statistical analysis.** Analyses were conducted on the intent-to-treat population, which included all randomized patients. Mean scores on the PROs over time in the stop and continuation group were compared using generalized estimating equation (GEE) models, with exchangeable correlation structures for repeated linear data with group (discontinuation versus continuation), time (0, 3, 6, and 12 months), and group by time interaction as categorical factors to detect any difference in PRO trajectories over time. Post hoc *t*-tests were performed to analyze between-group differences in scores at the different time points. Based on estimated marginal means and corresponding SEs from the GEE models, between-group standardized effect sizes were expressed as Cohen's d (Δ estimated marginal mean / pooled SD) with 95% confidence intervals (95% CIs), where 0.20 denotes a small, 0.50 a moderate, and 0.80 a large effect (41). Besides comparing mean scores on a group level, individual patient-level analyses were performed by calculating the proportion of patients experiencing a clinically important worsening of symptoms on each PRO at each time point as compared to baseline. Proportions of patients experiencing a worsening of symptoms greater or equal to the MCID were compared between the stop and continuation group using binary logistic GEE models, again with group, time (3, 6, and 12 months), and group by time as factors. Post hoc chi-square tests and relative risks (RRs) with 95% CIs were used to examine the significance and magnitude of differences at each followup point. Additional descriptive analyses (linear GEE models without adjustment for covariates) were performed in the stop group only to compare longitudinal scores on each PRO between patients who needed to restart TNFi within 6 months versus those that did not restart during the full 12 months of the study.

## RESULTS

**Patient characteristics and baseline scores.** Baseline demographics, disease characteristics, and PRO scores were similar in both groups (Table 1). Most patients had longstanding, RF-positive disease. The majority (85.9%) was on their first TNFi, predominantly adalimumab or etanercept. Scores on all PROs indicated relatively mild disease impact at baseline, as would be expected for RA patients in stable remission or LDA. Mean SF-36 component summary scores indicated that patients' mental health was comparable to general population norms, while their physical health score was approximately half an SD below the population mean. PRO scores were significantly but modestly correlated with DAS28-ESR scores over the different time points (Pearson's r ranging from 0.08 to 0.45), confirming that patient-perceived symptoms and clinical disease activity reflected related but distinct aspects of the disease (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract).

**Table 1. Baseline characteristics of the patients***

| Characteristic | Stop TNFi (n = 531) | Continue TNFi (n = 286) |
|---|---|---|
| Female, no. (%) | 362 (68.2) | 188 (66.0) |
| Age, mean ± SD years | 60.0 ± 11.8 | 59.7 ± 10.6 |
| Disease duration, mean ± SD years | 12.0 ± 8.8 | 11.1 ± 8.4 |
| DAS28, mean ± SD | 1.97 ± 0.76 | 2.05 ± 0.74 |
| RF positive, no. (%) | 328 (67.5) | 178 (67.4) |
| TNFi, no. (%) | | |
| Adalimumab | 271 (51.1) | 129 (45.1) |
| Etanercept | 213 (40.2) | 133 (46.5) |
| Infliximab | 25 (4.7) | 14 (4.9) |
| Golimumab | 15 (2.8) | 8 (2.8) |
| Certolizumab | 6 (1.1) | 2 (0.7) |
| TNFi taken, no. (%) | | |
| First | 459 (86.6) | 243 (85.0) |
| Second | 61 (11.5) | 37 (12.9) |
| Third | 10 (1.9) | 6 (2.1) |
| SF-36, mean ± SD | | |
| PCS score | 45.6 ± 8.8 | 45.3 ± 8.8 |
| MCS score | 52.0 ± 8.9 | 51.64 ± 10.2 |
| Bodily pain score | 71.9 ± 19.4 | 72.1 ± 19.3 |
| EQ-5D, mean ± SD | 0.84 ± 0.18 | 0.85 ± 0.14 |
| HAQ DI, mean ± SD | 0.60 ± 0.62 | 0.59 ± 0.59 |
| BRAF-MDQ, mean ± SD | 14.7 ± 11.2 | 15.4 ± 12.2 |

* TNFi = tumor necrosis factor inhibitor; DAS28 = Disease Activity Score in 28 joints; RF = rheumatoid factor; SF-36 = 36-Item Short Form health survey; PCS = physical component summary; MCS = mental component summary; EQ-5D = EuroQol 5-domain; HAQ DI = Health Assessment Questionnaire disability index; BRAF-MDQ = Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire.

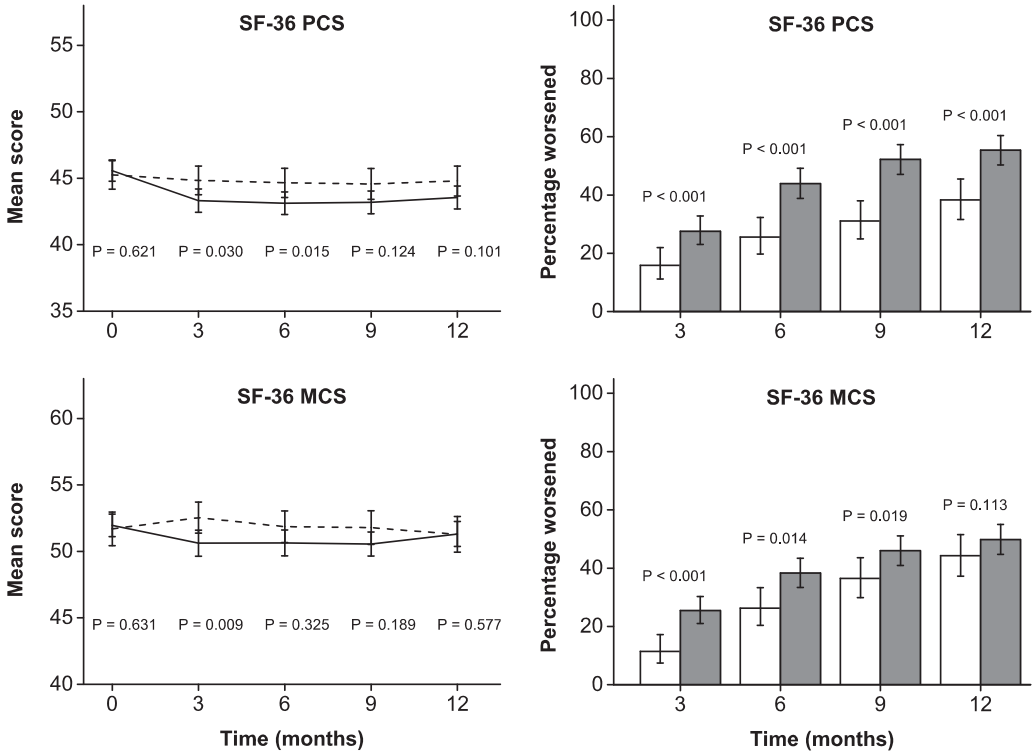


**Figure 1.** Estimated mean patient-reported outcome scores on physical and mental health status in all patients over time (left panels), stratified into those who continued (broken line) and stopped (solid line) tumor necrosis factor inhibitors. Cumulative proportion of patients experiencing a minimum clinically important difference as compared with baseline (right panels); gray bars represent patients in the stop group. Error bars are 95% Wald confidence intervals for both mean scores and proportions. *P* values are for between-group *t*-tests or chi-square tests. SF-36 = 36-Item Short Form health survey; PCS = physical component summary; MCS = mental component summary.

**PRO scores in patients stopping versus continuing TNFi.** The GEE analyses (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract) showed a significant group by time interaction for all PROs, except for the EQ-5D, indicating that mean scores changed significantly between both groups over time. Although all effect sizes were of small magnitude, post hoc *t*-tests confirmed that mean scores at 3 months were significantly worse in the stop group than those in the continuation group on all PROs (Figures 1, 2, and 3). Mean SF-36 MCS scores appeared to improve slightly between baseline and 3-month followup in the continuation group, but this improvement was not significant (paired *t*-test, $P = 0.192$).

The short-term impact of stopping TNFi was largest for pain with SF-36 BP scores being around 7 points lower (effect size −0.30) in the stop group at 3 months. After this point, PRO scores in the stop group tended to stabilize and steadily improve again. At 6 months, physical health status, pain, and functional disability scores remained significantly worse in the stop group. Pain in the stop group improved further and was no longer significantly different at 12 months. Functional disability scores, however, remained slightly (effect size 0.18) but significantly higher in the stop group.

**MCIDs in patients stopping versus continuing TNFi.** The proportion of patients experiencing a worsening greater or equal to the MCID at 3 months was significantly larger in the stop group for all PROs (Supplementary Table 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract) (Figures 1, 2, and 3). Except for the SF-36 MCS, there was no interaction between group and time, indicating that this difference between groups remained stable over the remaining course of followup. A substantial number of patients continuing TNFi also experienced an MCID due to natural fluctuations or measurement error. The proportion of patients experiencing an MCID was largest for pain at all followup time points in both the continuation (31.3% to 64.0%) and the stop group (49.5% to 80.0%). The RR for experiencing an MCID at 3 months ranged from 1.63 on the EQ-5D to 2.19 on the SF-36 MCS for patients stopping versus continuing TNFi. Over the course of followup, RRs decreased for all PROs, but the differences remained significant for all PROs, except mental health status (Figures 1, 2, and 3). After 12 months, trying to stop TNFi had resulted in an additional 5.5% (SF-36 MCS) to 18.6% (HAQ DI) of patients experiencing a clinically important worsening on top of those observed in the continuation group.

**PRO scores in patients restarting versus not restarting TNFi.** At 6 months, TNFi was restarted in 211 patients in the stop group (39.7%), whereas 279 patients (52.5%) did not restart TNFi during the entire 12 months of followup. Estimated mean scores in both subgroups are shown in

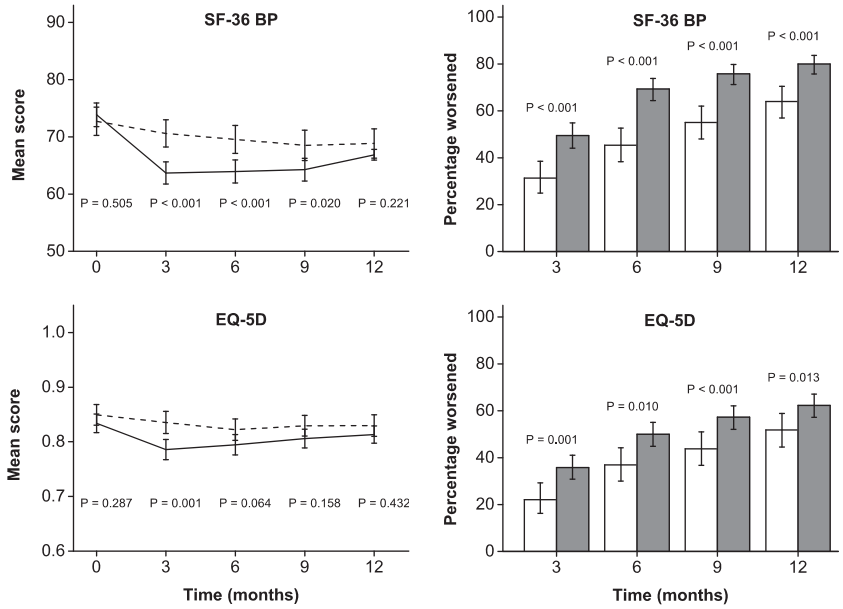


**Figure 2.** Estimated mean patient-reported outcome scores on physical and mental health status in all patients over time (left panels), stratified into those who continued (broken line) and stopped (solid line) tumor necrosis factor inhibitors. Cumulative proportion of patients experiencing a minimum clinically important difference as compared with baseline (right panels); gray bars represent patients in the stop group. Error bars are 95% Wald confidence intervals for both mean scores and proportions. *P* values are for between-group *t*-tests or chi-square tests. SF-36 = 36-Item Short Form health survey; BP = bodily pain; EQ-5D = EuroQol 5-domain.

Figures 4 and 5. In those patients experiencing a flare and needing to restart TNFi within the first 6 months after baseline, this was accompanied by a sharp worsening of physical health status, pain, health utility, and disability scores at 3 and 6 months. Effect sizes were −0.42 (95% CI −0.60, −0.24) and −0.42 (95% CI −0.60, −0.24) for SF-36 PCS, −0.56 (95% CI −0.74, −0.37) and −0.41 (95% CI −0.59, −0.22) for SF-36 BP, −0.36 (95% CI −0.54, −0.18) and −0.21 (95% CI −0.39, −0.03) for EQ-5D, and 0.30 (95% CI 0.12, 0.48) and 0.29 (95% CI 0.11, 0.47) for HAQ DI scores

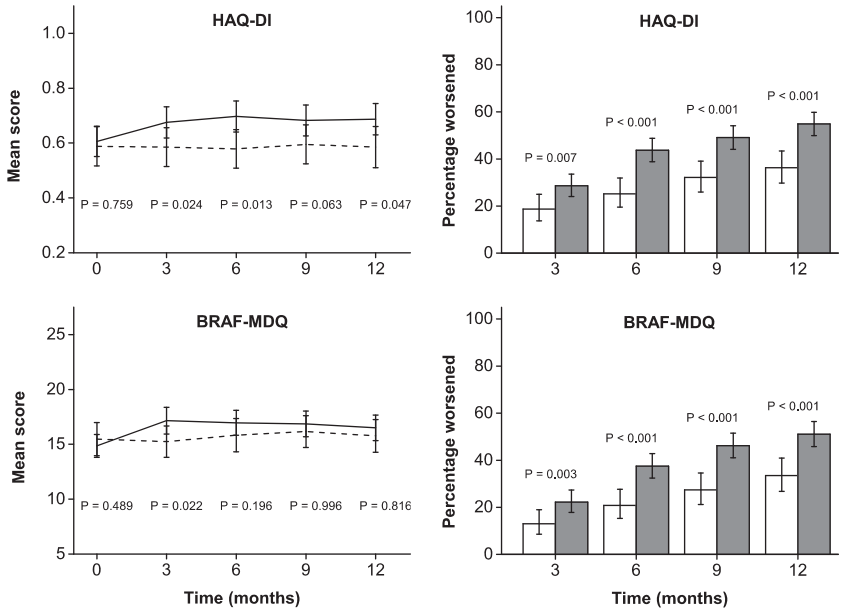


**Figure 3.** Estimated mean patient-reported outcome scores on physical and mental health status in all patients over time (left panels), stratified into those who continued (broken line) and stopped (solid line) tumor necrosis factor inhibitors. Cumulative proportion of patients experiencing a minimum clinically important difference as compared with baseline (right panels); gray bars represent patients in the stop group. Error bars are 95% Wald confidence intervals for both mean scores and proportions. *P* values are for between-group *t*-tests or chi-square tests. HAQ-DI = Health Assessment Questionnaire disability index; BRAF-MDQ = Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire.

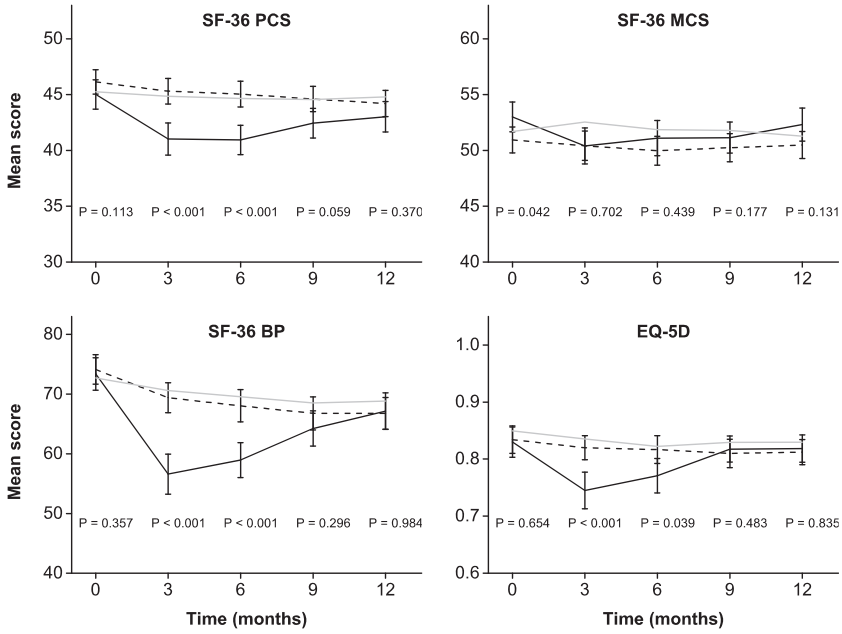


**Figure 4.** Estimated mean patient-reported outcome scores on physical and mental health status, pain, and health utility of patients stopping tumor necrosis factor inhibitors (TNFi), stratified into those who restarted within 6 months (solid black line; n = 211) and remained off TNFi throughout the 12-month study period (broken line; n = 279). Those who restarted after the 6-month visit were excluded. For comparison, mean scores in the total continuation group are plotted as solid gray lines (without error bars). Error bars are 95% Wald confidence intervals. *P* values are for between-group *t*-tests or chi-square tests. SF-36 = 36-Item Short Form health survey, version 2; PCS = physical component summary; MCS = mental component summary; BP = bodily pain; EQ-5D = EuroQol 5-domain.

at 3 and 6 months, respectively. PRO scores improved rapidly after restart and were no longer significantly different from those that did not restart TNFi at all after 9 months.

**Patient-reported experience of stopping TNFi.** The proportion of patients who rated their experience of stopping TNFi as positive or very positive after 12 months was 38.8%, while 43.6% experienced stopping as negative or very negative (see Supplementary Table 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract). At each time point, experience ratings were strongly associated with having restarted TNFi at that time or not. Across study visits, 72.4% to 76.2% of the patients that restarted TNFi rated their experience as negative, as opposed to 12.1% to 18.9% of the patients that did not restart (see Supple-

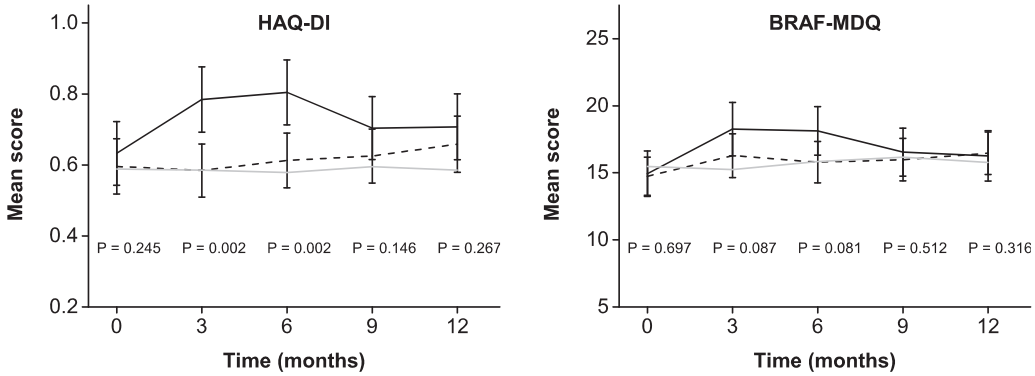


**Figure 5.** Estimated mean patient-reported outcome scores on disability and fatigue of patients stopping tumor necrosis factor inhibitors (TNFi), stratified into those who restarted within 6 months (solid black line; n = 211) and remained off TNFi throughout the 12-month study period (broken line; n = 279). Those who restarted after the 6-month visit were excluded. For comparison, mean scores in the total continuation group are plotted as solid gray lines (without error bars). Error bars are 95% Wald confidence intervals. *P* values are for between-group *t*-tests or chi-square tests. HAQ-DI = Health Assessment Questionnaire disability index; BRAF-MDQ = Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire.

mentary Table 5, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract).

## DISCUSSION

The POET trial showed that stopping TNFi treatment in patients with established RA in remission or with stable LDA resulted in substantially more clinical flares, but that most patients who restarted TNFi treatment quickly regained remission or LDA (16). The current study extends these findings by demonstrating that stopping TNFi also had a significant, but small, short-term negative impact on patient-reported physical and mental health status, health utility, pain, disability, and fatigue. Except for physical disability, PRO scores did not remain significantly different at 12 months. However, the proportion of patients with worsening above the MCID was higher and differed significantly for more PROs (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23315/abstract). Moreover, subgroup analyses showed that for patients who restarted TNFi after a flare within the first 6 months, 9- and 12-month PRO scores were no longer different from those that did not restart at all. As such, a strategy of trying to stop TNFi treatment and quickly restarting in case of flare does not appear to have substantial long-term consequences for the burden of disease as experienced by patients. To date, very few studies have thoroughly examined the longitudinal impact of stopping TNFi on patient-reported health status and symptoms. The double-blind PRESERVE study showed that after 52 weeks patients maintaining 50 mg etanercept plus methotrexate treatment had significantly better scores than patients using placebo plus methotrexate on general health, pain, disability, health utility, sleep, and fatigue (14). Some other nonrandomized or observational studies specifically reported the effects of stopping TNFi on 1 or more PRO, usually by comparing long-term disability scores in patients who did and did not successfully discontinue TNFi. The observational RRR (Remission Induction by Remicade in RA) study showed that 1-year mean HAQ DI scores were significantly lower in patients who remained in DAS remission after stopping infliximab versus those that did not (42). Post hoc analyses of 104 patients in the BeSt study (Treatment Strategies for Rheumatoid Arthritis) who discontinued infliximab after achieving stable LDA showed that HAQ DI scores at 1 and 3 years after discontinuation were similar to scores at discontinuation in both restarters and patients with sustained LDA (43).

However, 5 years after discontinuation, median disability scores were significantly increased in restarters. The open-label, nonrandomized HONOR study (Humira Discontinuation Without Functional and Radiographic Damage Progression Following Sustained Remission) did not show a significant difference in 1-year HAQ DI scores between patients that did and did not sustain remission after adalimumab discontinuation (44). Another very small observational study in 21 RA patients did not find significant differences in 1-year HAQ DI scores either (45), although with only 21 analyzed patients this study was likely underpowered for this purpose.

The POET study is the first large pragmatic trial examining the impact of trying to stop TNFi on both clinical and PROs in realistic clinical settings. Overall, the current findings show that stopping TNFi resulted in significantly worse scores on all PRO domains, particularly reflecting pain. However, the magnitude of differences was generally small and decreased again after 3–6 months. The proportion of patients experiencing a worsening greater or equal to the MCID within 3 months was also significantly larger in the stop group and remained significantly different at 12 months for most outcomes. At 12 months, the proportion of patients in the stop group that experienced an MCID was >50% on all PROs except mental health status. However, a substantial proportion of patients who continued TNFi also experienced an MCID, due to either natural disease fluctuations or measurement error. Like the effect sizes for mean scores, RRs for experiencing an MCID steadily decreased for all PROs over time.

In interpreting these findings, it is important to keep in mind that the POET study was an open-label pragmatic trial in which, after randomization, rheumatologists were free to prescribe and adjust medications as considered clinically necessary. Consequently, it is not possible to attribute the differences between groups purely to the effects of stopping TNFi, and the impact on PRO scores may be substantially confounded by other treatment decisions. For instance, rheumatologists are likely to have started or increased csDMARDs or nonsteroidal antirheumatic drugs (NSAIDs) in patients in the stop group, with an increase in disease activity but not meeting the DAS28 flare criterion, resulting in an underestimation of impact. Moreover, as the followup period of POET was limited to 1 year, it is unclear how many patients experienced a disease relapse after this period. Finally, the POET study only examined the effects of completely stopping TNFi. Recently, several studies have suggested that TNFi dose reduction may be more effective in maintaining LDA (14,15). This may also translate into less impact on patient-reported symptoms. For instance, a recent trial on disease activity–guided tapering of adalimumab or etanercept showed that HAQ DI and EQ-5D scores remained stable in the dose reduction group and did not differ significantly over time from patients continuing TNFi (46).

An important strength of the current study is that it included a broad range of PROs, covering all PRO domains as defined by the ACR and World Health Organization/International League of Associations for Rheumatology core set end points for RA clinical trials (47,48), as well as fatigue, which was more recently endorsed as a symptom important to patients (49). Moreover, well-established and validated measures were used for all domains. For instance, we used the multi-item BRAF-MDQ to measure fatigue, which has better content validity and measurement precision than more commonly used single-item measures of fatigue (40).

In summary, a strategy of trying to stop TNFi treatment in patients with stable LDA in daily clinical practice has a negative short-term impact on patient-perceived symptoms but does not appear to have substantial longer-term consequences for the burden of disease as experienced by patients.

## AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Ghiti Moghadam had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study conception and design.** Vonkeman, van Riel, van de Laar, Jansen.

**Acquisition of data.** Ghiti Moghadam, ten Klooster, Vonkeman, van Riel, van de Laar, Jansen.

**Analysis and interpretation of data.** Ghiti Moghadam, ten Klooster, Kneepkens, Klaasen, Stolk, Tchetverikov, Vreugdenhil, van Woerkom, Goekoop-Ruiterman, Landewé.

## REFERENCES


1. Hochberg MC, Tracy JK, Hawkins-Holt M, Flores RH. Comparison of the efficacy of the tumour necrosis factor α blocking agents adalimumab, etanercept, and infliximab when added to methotrexate in patients with active rheumatoid arthritis. Ann Rheum Dis 2003;Suppl 2:ii13–6.
2. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technol Assess 2006;10:1–229.
3. Alonso-Ruiz A, Pijoan JI, Ansuategui E, Urkaregi A, Calabozo M, Quintana A. Tumor necrosis factor α drugs in rheumatoid arthritis: systematic review and metaanalysis of efficacy and safety. BMC Musculoskelet Disord 2008;9:52.
4. Maini RN, Breedveld FC, Kalden JR, Smolen JS, Furst D, Weisman MH, et al. Sustained improvement over two years in physical function, structural damage, and signs and symptoms among patients with rheumatoid arthritis treated with infliximab and methotrexate. Arthritis Rheum 2004;50:1051–65.
5. Kimel M, Cifaldi M, Chen N, Revicki D. Adalimumab plus methotrexate improved SF-36 scores and reduced the effect of rheumatoid arthritis (RA) on work activity for patients with early RA. J Rheumatol 2008;35:206–15.
6. Kekow J, Moots RJ, Emery P, Durez P, Koenig A, Singh A, et al. Patient-reported outcomes improve with etanercept plus methotrexate in active early rheumatoid arthritis and the improvement is strongly associated with remission: the COMET trial. Ann Rheum Dis 2010;69:222–5.
7. Strand V, Smolen JS, van Vollenhoven RF, Mease P, Burmester GR, Hiepe F, et al. Certolizumab pegol plus methotrexate provides broad relief from the burden of rheumatoid arthritis: analysis of patient-reported outcomes from the RAPID 2 trial. Ann Rheum Dis 2011;70:996–1002.
8. Genovese MC, Han C, Keystone EC, Hsia EC, Buchanan J, Gathany T, et al. Effect of golimumab on patient-reported outcomes in rheumatoid arthritis: results from the GO-FORWARD study. J Rheumatol 2012;39:1185–91.
9. Dreyer L, Mellemkjær L, Andersen AR, Bennett P, Poulsen UE, Juulsgaard Ellingsen T, et al. Incidences of overall and site specific cancers in TNFα inhibitor treated patients with rheumatoid arthritis and other arthritides: a follow-up study from the DANBIO Registry. Ann Rheum Dis 2013;72:79–82.
10. Raaschou P, Simard JF, Holmqvist M, Askling J. Rheumatoid arthritis, anti-tumour necrosis factor therapy, and risk of malignant melanoma: nationwide population based prospective cohort study from Sweden. BMJ 2013;346:1939.
11. Ramiro S, Gaujoux-Viala C, Nam JL, Smolen JS, Buch M, Gossec L, et al. Safety of synthetic and biological DMARDs: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. Ann Rheum Dis 2014;73:529–35.
12. Tanaka Y. Intensive treatment and treatment holiday of TNF-inhibitors in rheumatoid arthritis. Curr Opin Rheumatol 2012;24:319–26.
13. Van Ingen IL, Lamers-Karnebeek F, Jansen TL. Optimizing the expediency of TNFi in rheumatoid arthritis: offering a TNFi holiday in patients having reached low-disease activity in the maintenance phase. Expert Opin Biol Ther 2014;14:1761–7.
14. 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;381:918–29.
15. Van Vollenhoven RF, Ostergaard M, Leirisalo-Repo M, Uhlig T, Jansson M, Larsson E, et al. Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. Ann Rheum Dis 2016;75:52–8.
16. 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 low disease activity: a pragmatic multicenter, open-label randomized controlled trial. Arthritis Rheumatol 2016;68:1810–7.
17. Hagglund KJ, Haley WE, Reveille JD, Alarcón GS. Predicting individual differences in pain and functional impairment among patients with rheumatoid arthritis. Arthritis Rheum 1989;32:851–8.
18. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis Care Res (Hoboken) 2010;62:647–56.
19. Studenic P, Radner H, Smolen JS, Aletaha D. Discrepancies between patients and physicians in their perceptions of rheumatoid arthritis disease activity. Arthritis Rheum 2012;64:2814–23.
20. Oude Voshaar AH, Taal E, ten Klooster PM, Vonkeman HE, van de Laar MA. Predictors of no improvement in subjective health perception in newly diagnosed rheumatoid arthritis patients with a good DAS28 response at 12 months [abstract]. Arthritis Rheum 2013;65:S921.
21. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66:40–1.
22. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
23. Van der Maas A, Lie E, Christensen R, Choy E, de Man YA, van Riel P, et al. Construct and criterion validity of several proposed DAS28-based rheumatoid arthritis flare criteria: an OMERACT cohort validation study. Ann Rheum Dis 2013;72:1800–5.
24. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight–joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.
25. Ware JE Jr. SF-36 health survey update. Spine (Phila Pa 1976) 2000;25:3130–9.
26. The EuroQol Group. EuroQo: a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
27. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.
28. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: a cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analog scales, and numerical rating scales. Arthritis Care Res (Hoboken) 2010;62:1559–68.
29. ten Klooster PM, Vonkeman HE, Taal E, Siemons L, Hendriks L, de Jong AJ, et al. Performance of the Dutch SF-36 version 2 as a measure of health-related quality of life in patients with rheumatoid arthritis. Health Qual Life Outcomes 2013;11:77.
30. Ware JE, Kosinski M, Dewey JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 health survey (standard & acute forms). Lincoln (RI): QualityMetric; 2000.

31. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000;43:1478–87.
32. Norman G, Sloan J, Wyrwich K. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582–92.
33. Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53–72.
34. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006; 15:1121–32.
35. Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, et al. A comparison of generic, indirect utility measures disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005;60:1571–82.
36. Peter M, ten Klooster PM, Taal E, van de Laar MA. Rasch analysis of the Dutch Health Assessment Questionnaire disability index and the Health Assessment Questionnaire II in patients with rheumatoid arthritis. Arthritis Rheum 2008; 59:1721–8.
37. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. J Rheumatol 2003;30:167–78.
38. Wolfe F, Pincus T. Listening to the patient: a practical guide to self-report questionnaires in clinical care. Arthritis Rheum 1999;42:1797–808.
39. Dures EK, Hewlett SE, Cramp F, Greenwood R, Nicklin JK, Urban M, et al. Reliability and sensitivity to change of the Bristol Rheumatoid Arthritis Fatigue scales. Rheumatology (Oxford) 2013;52:1832–9.
40. Oude Voshaar MA, ten Klooster PM, Bode C, Vonkeman HE, Glas CA, Jansen T, et al. Assessment of fatigue in rheumatoid arthritis: a psychometric comparison of single-item, multi-item, and multidimensional measures. J Rheumatol 2015;42: 413–20.
41. Cohen J. Statistical power analysis for the behavioural sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
42. Tanaka Y, Takeuchi T, Mimori T, Saito K, Nawata M, Kameda H, et al. Discontinuation of infliximab after attaining low disease activity in patients with rheumatoid arthritis: RRR (remission induction by Remicade in RA) study. Ann Rheum Dis 2010;69:1286–91.
43. Van den Broek M, Klarenbeek NB, Dirven L, van Schaardenburg D, Hulsmans HM, Kerstens PJ, et al. Discontinuation of infliximab and potential predictors of persistent low disease activity in patients with early rheumatoid arthritis and disease activity score-steered therapy: subanalysis of the BeSt study. Ann Rheum Dis 2011;70:1389–94.
44. Tanaka Y, Hirata S, Kubo S, Fukuyo S, Hanami K, Sawamukai N, et al. Discontinuation of adalimumab after achieving remission in patients with established rheumatoid arthritis: 1-year outcome of the HONOR study. Ann Rheum Dis 2015; 74:389–95.
45. Brocq O, Millasseau E, Albert C, Grisot C, Flory P, Roux CH, et al. Effect of discontinuing TNFα antagonist therapy in patients with remission of rheumatoid arthritis. Joint Bone Spine 2009;76:350–5.
46. Van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, et al. Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. BMJ 2015;350:1389.
47. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729–40.
48. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol 1994;21: 86–9.
49. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174–7.