RHEUMATOLOGY

Concise report

doi:10.1093/rheumatology/kex184

Limited value for ultrasonography in predicting flare in rheumatoid arthritis patients with low disease activity stopping TNF inhibitors

Femke B. Lamers-Karnebeek[1], Jolanda J. Luime[2], David F. Ten Cate[2], Steven Teerenstra[3], Nanno W. A. A. Swen[4], Andreas H. Gerards[5], Jos Hendrikx[6], Emma M. van Rooyen[1], Ramon Voorneman[7], Cees Haagsma[8], Natalja Basoski[9], Mike de Jager[10], Marjan Ghiti Moghadam[11], Monique N. Efde[12], Yvonne P. M. Goekoop-Ruiterman[13], Piet L. C. M. van Riel[6], Johannes W. G. Jacobs[14] and Tim L. Jansen[12]

Abstract

Objective. Ultrasonography (US) can be used for treatment decisions in RA patients. This study investigated the added value of US to clinical variables in predicting flare in RA patients with longstanding low disease activity when stopping TNF inhibitors (TNFi).

Methods. Cox models with and without using US added to clinical variables were developed in the Potential Optimization of Expediency of TNFi–UltraSonography study. RA patients (n = 259), using >1 year TNFi and csDMARD with DAS28 < 3.2 for 6 months prior to inclusion, were followed for 52 weeks after stopping TNFi. The added value of US was assessed in two ways: first, by the extent to which individual predictions for flare at 52 weeks with and without US differed; and second, by comparing how US information improved the prediction to classify patients at 52 weeks in the low risk (<33% flare), intermediate risk (33–50%) and high risk (50–100%) groups.

Results. Although US was predictive of flare at group level (multivariate hazard ratio = 1.7; 95% CI: 1.1, 2.5), individual predictions for flare at 52 weeks with and without US differed little (median difference 3.7%; interquartile range: −7.8 to 6.5%). With US, 15.9% of patients were designated low risk; without US, 14.6%. In fact, 12.0% of patients were US-classified as low risk with/without knowing US.

Conclusion. In RA patients with longstanding low disease activity, at time of stopping TNFi, US is a predictor for flare at group level, but at the patient level, US has limited added value when common clinical parameters are used already, though the predictive value of clinical predictors is modest as well.

Key words: rheumatoid arthritis, ultrasonography, remission, low disease activity, TNFi, prediction

CLINICAL SCIENCE

[1]Department of Rheumatology, Radboud University Medical Center, Nijmegen, [2]Department of Rheumatology, Erasmus Medical Center, Rotterdam, [3]Radboud Institute for Health Sciences, Department for Health Evidence, Radboud University Medical Center, Nijmegen, [4]Department of Rheumatology, Alkmaar Medical Center, Alkmaar, [5]Department of Rheumatology, Franciscus Gasthuis & Vlietland, Schiedam, [6]Radboud Institute for Health Sciences, IQ healthcare, Radboud University Medical Center, Nijmegen, [7]Department of Rheumatology, Reade Medical Centre, Amsterdam, [8]Department of Rheumatology and Clinical Immunology, Ziekenhuisgroep Twente, Almelo, [9]Department of Rheumatology, Maasstad Hospital, Rotterdam, [10]Department of Rheumatology, Albert Schweitzer Hospital, Dordrecht, [11]Department of Rheumatology, Arthritis Center Twente MST & University of Twente, Enschede, [12]Department of Rheumatology, Viecuri Medicall Center, Venlo, [13]Deparment of Rheumatology, HagaZiekenhuis, The Hague and [14]Department of Rheumatology & Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands

Submitted 24 November 2016; revised version accepted 6 April 2017

*Correspondence to: Femke B. Lamers-Karnebeek, Department of Rheumatology Bernhoven, Post Office box 707, 5406PT, Uden, The Netherlands.
E-mail: f.lamers@bernhoven.nl

Rheumatology key messages

- Ultrasonography is a modest flare predictor at group level in RA patients with DAS28 < 3.2 stopping TNFi.
- Ultrasonography has limited added value upon clinical parameters in RA patients with DAS28 < 3.2 stopping TNFi.

Introduction

The primary goal of treating patients with RA is to maximize long-term health-related quality of life through control of symptoms, prevention of structural damage, normalization of function and participation in social and work-related activities [1]. Since the use of treat-to-target and tight-control strategies [1], and the availability of biologic DMARDs, remission is more often achieved and maintained [2]. Biologics are very efficacious drugs to suppress disease activity, but they also give rise to an increased risk of infections and, though small, of skin malignancy [3]; in addition the cost of these agents are significant [4].

Based on these concerns, discontinuing biologics appears a viable approach for patients with sustained remission or low disease activity (LDA) [5, 6], but then there is a risk of recurrence of disease activity. However, in most patients with a disease flare a quick restart is effective again [7, 8].

The ideal profile of the patient who profits most from de-escalation or even stopping DMARD treatment remains to be defined [6]. Longer disease duration, smoking [9], higher BMI [10], longer use of TNFi [9, 11], higher DAS28 (ESR [11] and CRP [12]), presence of joint erosions [9], RF (IgM-RF) positivity [13] and anti-CCP [12], and high biomarker levels (multi-biomarker disease activity score) and MMP3 [13]) are markers of a less favourable disease course and potential predictors for flare. Also joint inflammation on imaging modalities such as MRI, PET and ultrasonography (US) [8, 14] is associated with increased risk of flare. Compared with MRI and PET, US is the most promising imaging tool to be studied for its capacity to predict flare, as it is generally readily available, cheaper and less time consuming.

Literature shows that US-detected residual synovitis is frequent and predicts at group level the risk of relapse and structural progression in RA patients with clinical remission (DAS < 1.6, DAS28-CRP/ESR < 2.3–2.6) [15]. Presence of power Doppler (PD) activity on US was the most accurate predictor for flare in RA patients in remission [15]. PD-detected synovitis may predict biologic therapy tapering failure in RA patients in sustained clinical remission [14]. So US might be promising in identifying RA patients who are at risk for flare when a biologic treatment is discontinued in case of clinical remission. It may be hypothesized that subclinical disease activity detected by US might predict flare.

This study, called Potential Optimization of Expediency of TNFi–UltraSonography (POET-US), was undertaken to investigate whether US at the time of stopping TNFi in RA patients with LDA, has added value in predicting flare to clinical data at group and individual levels.

Methods

Participants and study design

Two hundred and fifty-nine RA patients randomized to TNFi cessation were included in this study, the POET-US, from March 2012 until March 2014. This US cohort study is part of the nationwide randomized controlled trial (RCT) POET. This RCT compared stopping TNFi to continuing TNFi among adult RA patients in LDA or remission who were on TNFi for at least 1 year with concomitant csDMARDs without changes in medication at least 6 months prior to inclusion [7]. LDA or remission was present for at least 6 months, defined as DAS28 < 3.2 or rheumatologists' assessment and CRP <10 mg/l. This study was approved separately form the POET study by the Commissie Mensgebonden Onderzoek Arnhem-Nijmegen and participants gave their written informed consent for this study according to the Declaration of Helsinki.

At baseline, every 3 months and in case of suspected flare patients were evaluated by the treating rheumatologist and/or rheumatology nurse. At these visits, DAS28-ESR was performed.

Clinical and laboratory covariates

Based on literature, the following prognostic parameters were collected: age, sex, length, weight, disease duration, DMARD use, RF and anti-CCP-antibody status and presence of erosions on X-ray of hands and feet.

US assessment

US was performed in 17 participating centres by 18 experienced ultrasonographers (16 rheumatologists, 1 radiologist and 1 rheumatologist in training), who were blinded for clinical data. US took place as soon as possible (at least within 2 weeks after inclusion in the group of patients who stopped TNFi. US was performed following the EULAR guidelines concerning patient position and scanning planes [14]. Joints were evaluated using a semi-quantitative scoring system: grey scale 0–3 (GS) and PD 0–3 using Naredo's modification of Skudlarek's scoring system [15]. PD was not performed if GS was 0. Twenty joints were scanned: bilateral MCP 1–5 (dorsal and ventral), the radio carpal and intercarpal joints of both wrists and bilateral MTP 2–5 (dorsal aspect). The US reliability was optimized by training and calibration sessions and using Naredo's modification of Skudlarek's scoring system [15]. US arthritis was defined as GS > 1 and/or PD > 0 based on literature [16].

Definition of flare

A flare was defined as a >0.6 point increase of DAS28 since study start and a DAS28 ⩾ 3.2. This definition is according to the proposition of an OMERACT study,

TABLE 1 Baseline patient's characteristics who stop TNFi treatment

Patient characteristic	No-flare (n = 114)	Flare (n = 112)	P-value
Age, mean (S.D.), years	58 (12)	61 (10)	0.097
Disease duration, median (IQR), years	8.0 (5–13)	11 (5–19)	0.058
Female, n (%)	75 (66)	74 (66)	0.964
BMI, mean (S.D.)	25.6 (4.81)	26.1 (3.68)	0.242
IgM RF positive, n (%)	77 (67.5)	86 (76.8)	0.121
Anti-CCP positive, n (%)	79 (69.3)	81 (72.3)	0.617
Erosive, n (%)	70 (61)	69 (62)	0.975
DAS28 at inclusion, mean (S.D.)	1.91 (0.88)	2.16 (0.68)	0.015
MTX, n (%)	100 (87.7)	97 (86.6)	
TNFi, n (%)			
Adalimumab	61 (48)	66 (52)	
Etanercept	51 (52.6)	46 (47.4)	
Infliximab	7 (53.8)	6 (46.2)	
US signs of arthritis $\geqslant$1 joint, n (%)	62 (54)	77 (69)	0.026

IQR: interquartile range.

validating DAS28-based flare criteria [17], and has also been used in the recently published POET paper [7].

Statistics analysis

Sample size

Literature suggests that 25–56% of patients who discontinue TNFi will have sustained biologic free remission if csDMARDs are continued. Using a flare rate of 50% in the first year and the common rule of thumb for including one predictor per 10 flares, at least 250 patients were necessary to answer our research question [18].

Patients who were lost to follow-up without flare were censored at last follow-up date. To assess the added value of US, a multivariate Cox model including the clinical and laboratory factors only and a model containing these and US data was fitted to the data, both models without covariate selection. The best functional form of and validity of the proportional hazard assumption of the variables in each model was determined using cumulative sums of martingale-based residuals [19]. Each model's fit was checked using deviance residuals plots and calibration plots. For the Cox prediction models the concordance index (c-index) was used to estimate the discriminative ability of each model as defined by Harrell [20]. Next, for each patient the probability of flare within 52 weeks as predicted from the model with and without US was compared. Also, (dis)agreement between the two models concerning classification into 0–33, 33–50 and >50% risk to flare within 52 weeks was assessed. The risk classification boundaries were motivated by the fact that even when TNFi is continued patients have a 1-year risk of relapse of 18.4% [7].

Results

Baseline characteristics

DAS28 at baseline was available in 252 out of 259 (97%) patients. Sixty-six per cent of all patients were female, the average age was 59 years, median disease duration was 9 years, and 72% RF and 71% ACPA positive (Table 1). Complete baseline data were available for 226 patients. Among the 226 patients flare was seen in 112 patients within 12 months with a median time-to-flare of 14 weeks (interquartile range (IQR) = 1–27). Baseline characteristics of non-completers (n = 26) were not significantly different from the complete cases.

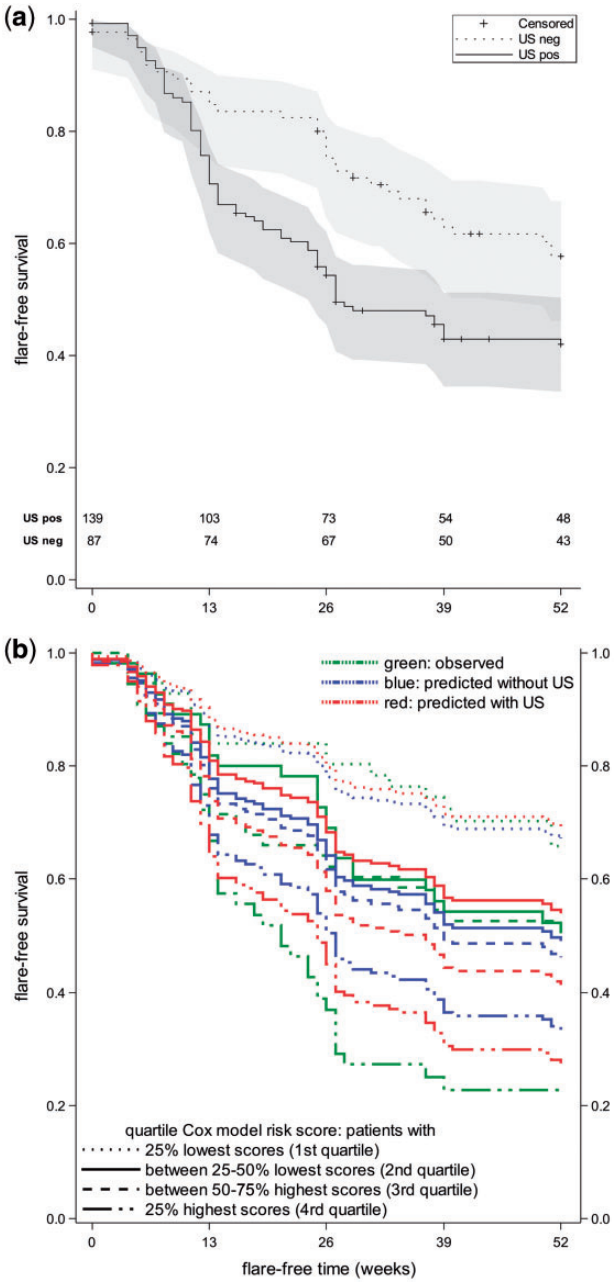
Comparison of predictions with and without US at 52 weeks

Having at least one joint with US inflammation resulted in a hazard ratio of 1.7 [Fig. 1A (95% CI: 1.1, 2.5)] for flare in the multivariate Cox model at 52 weeks. Fit of both models to the data was reasonably good as can be seen from the agreement between observed and predicted flare-free survival curves (see Fig. 1B). The c-index provides the probability that for any two patients A and B, of whom A has longer predicted time to flare than B based on the model, A had a longer actual time to flare than B. The discriminative ability of both models was comparable and modest: c-index = 0.62 (95% CI: 0.56, 0.68) without and c-index = 0.64 (95% CI: 0.59, 0.70) with US. Comparing individual predictions of flare at 52 weeks gave a median difference of 3.7% (IQR = −7.8 to 6.6%) between using US in the prediction model *vs* not including US. When classifying the individual predictions at 52 weeks into low risk (<33% flare risk), intermediate risk ($\geqslant$33–50%) and high risk (>50%), the overall agreement between the two models was 73.9%. Prediction with US designated 15.9% (36/226) patients as low risk compared with 14.6% (33/226) when no US was used; 27 patients were classified low risk by both models. For the intermediate risk group, these numbers were 27.9% *vs* 25.7% and in the highest group 56.2% *vs* 59.7%.

Discussion

In this 12-month prospective study 114 (50.4%) of the patients were able to stop TNFi without a flare in the first year. Based on their clinical characteristics we were

Fig. 1 Ultrasonography-findings at baseline and calibration plot illustrating fit of Cox-model



(**A**) Kaplan–Meier curve for non-US-arthritis at baseline *vs* US-arthritis at baseline (hazard ratio = 1.7; 95% CI: 1.1, 2.5). (**B**) Calibration plot illustrating the fit of the Cox model with US (red) and without US (blue) to the observed survival (green) in quartile groups based on the Cox score from the model with US.

able to identify three groups: those with low, medium and high risk of flare. Adding US to these clinical characteristics was informative at group level (multivariate hazard ratio for flare 1.7; 95% CI: 1.1, 2.5). However, when classifying patients in the low, intermediate and high risk groups for deciding to stop TNFi, this effect disappeared. It was anticipated that US would be more sensitive to reveal residual inflammatory activity that would predict a flare within the first months after cessation of TNFi given earlier publications [8, 14]. We could not affirm this.

Residual synovitis on US is common in 25–55% of the patients in LDA or remission [6]. Previously, associations were found for the risk of flare and structural progression in RA patients [15]. Also presence of US-arthritis was found to be a predictor for flare within a short term [6–12 months) after discontinuation of biologic agents

[14]. All studies evaluated US as a predictor at group level similar to our study. However, up to now, no data was provided at patient level necessary for the management of individual patients. This present study shows that US has little value as a predictor of flare in the individual patient when added to clinical data.

We conducted a prospective study design using RA resembling our daily rheumatology practice, thus allowing for generalizability. The follow up time of one year was sufficiently long in order to capture disease flares related to stop of TNFi. A limitation of the study may be that, since we selected patients in LDA or remission, little residual US inflammation was detected, which may have limited the predictive ability of US. Also, some clinical predictors were not collected in the POET(-US) study such as smoking, duration of TNFi use and duration of DMARD use before start of TNFi.

A final limitation could be that the model was developed on complete cases (n = 226), instead of using imputation on the full dataset (n = 259). Since only a small number of patients had data missing (13%), and moreover these patients had no statistically significant or clinically relevant difference from patients having complete baseline characteristics, this is not expected to change the developed models.

In conclusion, US at patient level at the moment of deciding to stop TNFi did not contribute to identifying the individual patients with a flare at a subsequent time point during the first year after TNFi cessation.

Acknowledgements

The authors wish to thank all patients, rheumatology nurses, participating rheumatologists (in training) and the radiologist of the participating centres.

Funding: This was an investigator initiated trial with an unrestricted grant of Abbvie.

Disclosure statement: P.L.C.M.vR. has received study grants, honoraria for consultancy and lecturing from Abbvie, Bristol-Myers Squibb, Merck Sharp & Dome, Pfizer, Roche, Eli Lilly, Galapagos, UCB-pharma and Celltrion. T.L.J. has received an unrestricted grant from Abbvie for additional POET projects including the Ultrasound Substudy (POET-US). All other authors have declared no conflicts of interest.

References

1 Smolen JS, Breedveld FC, Burmester GR *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2016;75:3–15.

2 Nam JL, Ramiro S, Gaujoux-Viala C *et al*. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis 2014;73:516–28.

3 Pereira R, Lago P, Faria R, Torres T. Safety of anti-TNF therapies in immune-mediated inflammatory diseases: focus on infections and malignancy. Drug Dev Res 2015;76:419–27.

4 Kievit W, Adang EM, Fransen J *et al*. The effectiveness and medication costs of three anti-tumour necrosis factor alpha agents in the treatment of rheumatoid arthritis from prospective clinical practice data. Ann Rheum Dis 2008;67:1229–34.

5 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;42:2012–22.

6 Schett G, Emery P, Tanaka Y *et al*. Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. Ann Rheum Dis 2016;75:1428–37.

7 Moghadam MG, Vonkeman HE, Ten Klooster PM *et al*. Stopping tumor necrosis factor-inhibitors in patients with established rheumatoid arthritis in remission or stable low disease activity: a pragmatic randomized multicenter open-label controlled trial. Arthritis Rheumatol 2016;68:1810–7.

8 Alivernini S, Peluso G, Fedele AL *et al*. Tapering and discontinuation of TNF-alpha blockers without disease relapse using ultrasonography as a tool to identify patients with rheumatoid arthritis in clinical and histological remission. Arthritis Res Ther 2016;18:39.

9 van den Broek M, Klarenbeek NB, Dirven L *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.

10 Kavanaugh A, Lee SJ, Curtis JR *et al*. Discontinuation of tumour necrosis factor inhibitors in patients with rheumatoid arthritis in low-disease activity: persistent benefits. Data from the Corrona registry. Ann Rheum Dis 2015;74:1150–5.

11 Tanaka Y, Hirata S, Kubo S *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.

12 Emery P, Burmester GR, Bykerk VP *et al*. Evaluating drug-free remission with abatacept in early rheumatoid arthritis: results from the phase 3b, multicentre, randomised, active-controlled AVERT study of 24 months, with a 12-month, double-blind treatment period. Ann Rheum Dis 2015;74:19–26.

13 Nishimoto N, Amano K, Hirabayashi Y *et al*. Drug free REmission/low disease activity after cessation of tocilizumab (Actemra) Monotherapy (DREAM) study. Mod Rheumatol 2014;24:17–25.

14 Naredo E, Valor L, De la Torre I *et al*. Predictive value of Doppler ultrasound-detected synovitis in relation to failed tapering of biologic therapy in patients with rheumatoid arthritis. Rheumatology 2015;54:1408–14.

15 Nguyen H, Ruyssen-Witrand A, Gandjbakhch F *et al*. Prevalence of ultrasound-detected residual synovitis and risk of relapse and structural progression in rheumatoid arthritis patients in clinical remission: a systematic review and meta-analysis. Rheumatology 2014;53:2110–8.

16 Witt M, Mueller F, Nigg A *et al*. Relevance of grade 1 gray-scale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. Arthritis Rheum 2013;65:1694–701.

17 van der Maas A, Lie E, Christensen R *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.

18 Steyerberg EW. Clinical Prediction Models: a Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.

19 Lin D, Wei LJ, Ying Z. Checking the cox model with cumulative sums of martingale-based residuals. Biometrika 1993;80:557–72.

20 Steyerberg EW, Vickers AJ, Cook NR *et al*. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.