Concise report

Ultrasound Doppler measurements predict success of treatment with anti-TNF-α drug in patients with rheumatoid arthritis: a prospective cohort study

Karen Ellegaard[1], Robin Christensen[1,2], Søren Torp-Pedersen[1],
Lene Terslev[1], Christian C. Holm[1], Merete J. Kønig[1], Peter S. Jensen[1],
Bente Danneskiold-Samsøe[1,3] and Henning Bliddal[1,3]

Abstract

Objective. To investigate the predictive ability of core outcomes applied in RA trials, including ultrasound (US) Doppler (USD) measurements differentiating patients who remain on anti-TNF-α therapy following 1 year.

Methods. Patients with RA in anti-TNF-α therapy were followed 1 year after therapy initiation. All patients had wrist involvement. At baseline, 2 weeks, 26 weeks and 1 year a USD examination, clinical examination including tender and swollen joint count, visual analogue scale (VAS) global and HAQ, biochemical measures and 28-joint DAS (DAS28) were collected for all patients. The amount of USD signal in the synovium was quantified by measuring the percentage of colour pixels—the colour fraction (CF). Predictive validity for patients who remain on anti-TNF-α therapy after 1 year was assessed for both USD measurements and other disease measures. Baseline values of disease measures of patients who remained on treatment after 1 year was compared with those who stopped therapy.

Results. The study cohort consisted of 109 patients. In this study, the baseline CF was the only measure predicting which patients would stay on the initial anti-TNF-α therapy for 1 year, evaluated using the square-root of CF ($P = 0.024$). The other disease markers could not significantly differentiate between the two groups of patients, with P-values of 0.86 and 0.98 for tender and swollen joint count, respectively, 0.86 for CRP, 0.24 for VAS, 0.10 for HAQ and 0.38 for DAS28.

Conclusion. There is now evidence to support that baseline USD, in contrast to clinical measures, can predict which patients will remain on anti-TNF-α 1 year after initiating therapy.

Key words: Rheumatoid arthritis, Ultrasound, Wrist, Inflammation, Synovium.

CLINICAL SCIENCE

Introduction

Anti-TNF-α agents are effective in many patients with RA [1, 2], while not effective in all [3], with predictors of response being necessary [4]. Ultrasound (US) Doppler (USD) has been suggested as a further predictor of disease activity [5]. Cross-sectional studies have shown concurrent validity between Doppler and other validated measures of disease activity [6–13]. Furthermore, USD has been applied in the assessment of patients with RA in treatment with anti-TNF-α, showing the ability of USD measurements to aid in the monitoring of treatment [14–21].

USD assessment of a single wrist joint has good concurrent validity when compared with validated measures of disease activity in RA [13]. Approximately 70% of patients with RA have affection of a wrist joint [22] and the wrist joint is therefore well-suited for estimation of arthritis activity by imaging [23].

Predictive validity refers to a measurement ability to forecast future events [24, 25]. Part of the assessment

recommended in the outcome measures in rheumatoid arthritis clinical trials (OMERACT) filter is to investigate the predictive value of a measurement.

The aim of this study was to investigate the ability of clinical measures and USD measurements in a wrist joint to predict which patients with RA will benefit from treatment with anti-TNF-α in terms of staying on therapy for at least 1 year. A second objective was to investigate the ability of USD measurements to monitor changes over time.

Methods

Patients

Patients with RA treated with an anti-TNF-α drug (adalimumab, etanercept or infliximab) were studied, all fulfilling the ACR criteria for RA [26]. The patients were enrolled consecutively in an outpatient clinic in 2003–07 as they started treatment. The patients were investigated at baseline and 1 year onwards with clinical examinations, blood tests and a USD examination of an individually defined target joint. The target joint was the joint with most pronounced synovial hyperaemia on USD; if possible, the most affected wrist. Only patients with wrist involvement were included in our data analysis. Both the radio-carpal and inter-carpal compartments of the wrist joint were included in the evaluation. Only the dorsal aspect of the wrist was evaluated to reduce time spent on the US examination. In our experience, the dorsal aspect of the joint is more frequently involved than the volar aspect. The local ethics committee (the committees on Biomedical Research Ethics for the Capital region of Denmark) approved the study (KF01-045/03) and informed consent was obtained.

Clinical and para-clinical examination

At each visit, the patients underwent assessment of the number of tender and swollen joints by a rheumatologist unaware of the results of the US examination. Blood was tested for ESR and CRP; the patients filled in an HAQ and a general health score on a visual analogue scale (VAS) global and a 28-joint DAS (DAS28)–CRP was calculated [27].

US examination

Scanning was performed with a US machine (Siemens, Mountainview, CA) using a linear array transducer with 14 MHz centre frequency. The Doppler pre-set was made according to the recommendations made by Torp-Pedersen and Terslev [28]. No adjustments of Doppler parameters were performed. The dorsal wrist was scanned longitudinally from side to side and images were obtained in the radial, central and ulnar positions. In the three positions, the area with the most pronounced Doppler activity was identified, the transducer was held in this position for a couple of heart cycles whereupon the image was frozen. The image with most Doppler activity was then selected and stored. This image acquisition technique has shown a test–retest reliability intraclass correlation coefficient (ICC) of 0.77 [29].

Four persons performed the US examinations (K.E., S.T.-P., L.T. and M.J.K.). One investigator (S.T.-P.) is the head of the US unit, with 20 years US experience; the other investigators had several years of US training and 1 month of specific training on the wrist scans.

Image analysis

USD activity was quantified by the colour fraction (CF) [6, 30]. The CF is the number of colour pixels divided by the total number of pixels in a region of interest (ROI). We defined ROI as the synovial tissue seen as a predominantly hypo-echoic mass located dorsally to the ulnar head and the radio-carpal and inter-carpal joints. In all three positions, the synovial tissue was traced and the CF was calculated (supplementary fig. 1, available as supplementary data at *Rheumatology* Online). Subsequently, the average of the three CF values was computed. In a previous study, this evaluation technique showed an excellent ICC [30].

Two investigators (K.E. and P.S.J.) with long-standing experience in image evaluation performed all image analyses blinded to the patients' clinical characteristics. The CF calculation was made in Datapro (DataPro; Noesis Courtaboeuf, France).

End points

It is an underlying assumption that if patients continue on a certain therapy for a period of time it is because there is both an effect of the drug and it is well tolerated. Thus, the continuation of a therapy can be used as a proxy for overall effectiveness and safety [31]. This assumption is supported by the fact that the main reason for patients with RA to discontinue a treatment is lack of efficacy [31]. Furthermore, correlation between withdrawal and ACR-20 response has been demonstrated [32].

Statistical analyses

In order to evaluate which, if any, of the variables measured at baseline could predict patients' survival on therapy for 1 year, we used Still-on-therapy as a dichotomization factor with two levels (yes/no). To evaluate the impact (i.e. clinical significance) of each of these possible predictive variables, we calculated the so-called effect size (ES) being the Cohen's index or the standardized mean difference—calculated as the mean difference between the groups (Still-on-therapy: yes *vs* no) divided by the S.D. in the total data set [33]; i.e. the higher the absolute value of the ES—the more predictive. Subsequently, based on the patients still on therapy after 1 year, we wanted to evaluate the discriminant capacity for each of the variables included [34]. This was evaluated using the standardized response mean (SRM).

Results

Patient characteristics

One hundred and sixty-two outpatients started on anti-TNF-α therapy; of these, 109 had wrist involvement.

At 1 year follow-up, 78 patients were still on therapy with the initial anti-TNF-α. These 78 were categorized as completers. The 31 patients who discontinued treatment were considered drop-outs. The reasons for drop-out were lack of efficacy (23 patients) or side effects (8 patients). Baseline data of the 109 patients are shown in Table 1. Predominantly females (71%) participated. The mean age was 58 years (range 26–84 years). The mean disease duration was 10.4 years (range 1–34.6 years), mean baseline CF was 0.24 (range 0.000–0.690) and mean DAS28 5.07 (range 1.40–7.80). When applying the EULAR DAS28 criteria for assessment of disease activity at baseline [34, 42] 4 had a DAS28 < 2.6 (2 completers and 2 drop-outs); 7 had < 3.2 (6 completers and 1 drop-out) and 33 had < 5.1 (19 completers and 14 drop-outs); the remaining 64 patients had a DAS28 > 5.1 (47 completers and 17 drop-outs). No statistically significant difference in concomitant therapies was seen between the completer and the drop-out groups (Table 1).

Completers *vs* drop-outs in measures of disease activity

The CF had a low efficiency and poor consistency due to a large statistic variation. Following assessment of multiple transformations, it was evident that a square-root transformation resulted in an unbiased estimate with improved efficiency and better consistency (data not shown). Thus, we use the square-root CF ($\sqrt{CF}$) as the primary US variable. As presented in Table 1, the baseline $\sqrt{CF}$ could predict those remaining on therapy ($P = 0.024$). Also, the CF was statistically significant ($P = 0.020$). No other statistically significant differences were found between the two groups. Furthermore, the square-root transformed CF had a moderate-to-large clinical predictive value, higher than any other measurements, in terms of differentiating patients who would still be on therapy after 1 year (ES = 0.57)—indicating that patients with low (or no) CF probably would withdraw from therapy.

In order to explore thresholds, an iterative *post hoc* analysis showed that the best cut-off for predicting those completing 1-year therapy were those with a $\sqrt{CF} > 0.23$ (i.e. CF ⩾ 0.05). As indicated in Table 1, the majority of the included patients had a CF value of at least 5% at baseline before initiating biologics. When using this threshold, most patients would remain on therapy (71.6%); of course, this threshold is to be considered preliminary, and need confirmation in a prospective trial.

At 1-year follow-up, USD data were only available for 69 patients, allowing us to only calculate a change score for these. Table 2 presents the discriminant capacity for the various outcome variables varied. It was evident that DAS28 was the most discriminative outcome measure when assessing changes following therapy—with an SRM of −1.45 (95% CI −1.83, −1.07)—having the best signal-to-noise ratio. In comparison, the SRM for $\sqrt{CF}$ was −0.62 (95% CI −0.91, −0.32) and CF was −0.52 (95% CI −0.81, −0.23).

Discussion

The main finding in this prospective cohort study was that USD measurement using CF obtained at baseline was the only outcome measure that could significantly predict which patients would remain on anti-TNF therapy. The ES for the CF was substantially higher than for any of the other disease activity measures [35]. CF gives the clinician a measure of hyperaemia, an integral part of inflammation, and the values are readily perceptible as high values of CF equal high inflammatory activity. It seems natural that an individual with high baseline inflammatory activity would benefit most from anti-inflammatory medication. We anticipated that DAS28 also could predict which patients would remain on therapy, but our results could not confirm this. This finding is in accordance with another study showing lack of predictive value of DAS28 [36]. Since lack of efficacy is the main reason for dropping out of therapy [31], this suggests that CF can be used as predictor of anti-TNF-α treatment success in patients with RA [31, 32].

Some overlap between the responders and non-responders was seen. This overlap naturally weakens the value of the CF as a predictive marker in a clinical setting; however, compared with the other markers of disease activity, the ES of CF was considerably higher. It can be argued that estimation of CF is too time consuming to be used in daily clinical praxis. However, software that allows estimation of the CF on the US machine is under preparation. This application will make it possible to estimate the CF at the US examination.

Some limitations of USD measurements exist when they are assessed according to the OMERACT filter [34]. There is very sparse information of US findings in healthy persons [37, 38]; differences in Doppler findings due to differences in machines or settings remain to be addressed [28] and furthermore, there is no consensus in the use of scoring systems. The perfusion in the synovial tissue is estimated from the part of the synovial area covered by colour. When given as a fraction, a theoretical bias is that a decrease in perfusion may be accompanied by a shrinking in synovial volume; this would result in an increase in the relative amount of colour, and a benefit of treatment might be overlooked, no matter which scoring system is applied. The only way to overcome the error caused by shrinking synovial tissue is to use an ROI defined by anatomical structures surrounding the synovial tissues [29].

The OMERACT filter emphasizes that the outcome may discriminate between groups. Thus, a good outcome measure needs to have a high signal-to-noise ratio. We found DAS28 to be a better discriminator than CF with an SRM of −1.45. The differentiation between best predictor (CF) and best discriminator (DAS28) is intriguing. The higher SRM of DAS28 as compared with CF may reflect that DAS28 overestimates the effect of therapy, and that the smaller decrease in CF indicates that the patients did only to some extent benefit from treatment. This interpretation is supported by a number of studies showing that patients with RA assessed clinically as being in remission

TABLE 1 Demographic, clinical and laboratory data of the 109 RA patients at baseline comparison of demographic, clinical and laboratory data at baseline between completers[a] and drop-outs[b]

Variable	Total sample (n = 109)	Completers (n = 78)	All drop-outs (n = 31)	Drop-outs (lack of efficacy) (n = 23)	P-value completers vs all drop-outs	ES	P-value completers vs drop-outs due to lack of efficacy	ES
Females	78 (71.1)	56 (71.8)	22 (71.0)	17 (74)	0.930[c]			
Age, years	57.9 (13.9) (25.5–84.3)	58.4 (13.3)	56.5 (15.5)	54.4 (16.3)	0.545[d]	0.14	0.293	0.28
Duration of RA	10.4 (9.0) (1.0–34.6)	9.7 (8.7)	12.0 (9.8)	11.8 (9.7)	0.355[e]	−0.25	0.346	−0.24
Height	170 (8) (148–191)	170 (9)	169 (7)	169 (7)	0.818[d]	0.05	0.536	0.13
Weight	72.8 (15.6) (44.0–130)	73.9 (16.6)	70.2 (12.5)	69.8 (11)	0.215[d]	0.24	0.173	0.26
BMI	25.3 (4.8) (16.9–45.0)	25.5 (4.9)	24.6 (4.6)	24.6 (4)	0.334[d]	0.20	0.356	0.20
DAS28	5.07 (1.36) (1.40–7.80)	5.14 (1.34)	4.88 (1.42)	5.1 (1.2)	0.375[d]	0.20	0.943	0.02
CRP	22 (28) (1–152)	23 (29)	20 (27)	24 (29)	0.453[e]	0.11	0.858	−0.04
TJC	11 (8) (0–28)	11 (8)	10 (10)	11 (9.5)	0.321[e]	0.14	0.861	0.05
SJC	8 (6) (0–26)	8 (6)	8 (8)	8 (7.4)	0.486[e]	0.01	0.975	0.01
Patient global VAS	63 (22) (6–100)	61 (23)	66 (20)	70 (19)	0.240[d]	−0.24	0.081	−0.38
HAQ-20	1.19 (0.63) (0.00–2.63)	1.25 (0.59)	1.03 (0.71)	1.1 (0.8)	0.098[e]	0.35	0.416	0.23
CF	0.243 (0.176) (0.000–0.690)	0.270 (0.180)	0.176 (0.151)	0.179 (0.151)	0.013[e]	0.53	0.020	0.51
$\sqrt{CF}$	0.450 (0.202) (0.000–0.831)	0.484 (0.188)	0.363 (0.215)	0.371 (0.206)	0.008[d]	0.60	0.024	0.57
MTX	49.5%	38 (50.6)	16 (59.2)		0.602			
Prednisolone	37.6%	28 (37.3)	13 (48.1)		0.779			
SSZ	10.1%	9 (12)	2 (7.4)		0.475			
Etanercept	30 (27.5)	25 (83.3)	5 (16.7)	4 (17.4)	0.110[f]			
Adalimumab	61 (56.0)	43 (70.5)	18 (29.5)	14 (60.8)	0.110[f]			
Infliximab	18 (16.5)	10 (55.6)	8 (44.4)	5 (21.7)	0.110[f]			

Continuous data are presented as mean (s.d.) and (range); counts are presented as the observed n (%). [a]Patients treated with the same anti-TNF-α drug for year. [b]Patients treated with the same anti-TNF-α drug for <1 year. [c]Analysed using a chi-square test with a factor for sex and a factor for completer. [d]Analysed using PROC TTEST, with unequal variances (i.e. the Satterthwaite method). [e]Analysed using PROC NPAR1WAY (Wilcoxon) (i.e. the Kruskal–Wallis test). [f]Analysed using a chi-square test with a factor for the label of anti-TNF applied and a factor for completer. ES calculated as the standardized mean difference; TJC: tender joint count; SJC: swollen joint count.

TABLE 2 Changes following 1 year of anti-TNF-α treatment

Variable	*n*	Mean (S.D.)	*t*-value	*P*-value	ρ	SRM (95% CI)
△CRP	75	−14.67 (23.90)	−5.31	<0.0001	0.532	−0.61 (−0.85, −0.37)
△DAS28	75	−2.16 (1.49)	−12.56	<0.0001	0.298	−1.45 (−1.83, −1.07)
△HAQ20	75	−0.43 (0.47)	−7.98	<0.0001	0.687	−0.92 (−1.14, −0.71)
△TJC	75	−8.21 (8.46)	−8.41	<0.0001	0.170	−0.97 (−1.32, −0.62)
△SJC	75	−5.97 (5.82)	−8.89	<0.0001	0.300	−1.03 (−1.36, −0.70)
△CF	69	−0.10 (0.19)	−4.31	<0.0001	0.318	−0.52 (−0.81, −0.23)
△$\sqrt{CF}$	69	−0.14 (0.23)	−5.13	<0.0001	0.339	−0.62 (−0.91, −0.32)
△VAS	75	−29.12 (30.11)	−8.38	<0.0001	0.138	−0.97 (−1.33, −0.61)

All change scores are calculated as post- and pre-treatment with anti-TNF-α. S.D.: S.D. of change scores; *t*-value: calculated from the mean change from baseline (i.e. one-sample *t*-test); ρ: the correlation between the pre- and post-scores. SRM: SRM from baseline; △$\sqrt{CF}$: $\sqrt{CF}_{post} - \sqrt{CF}_{pre}$; TJC: tender joint count; SJC: swollen joint count.

still had Doppler activity [5, 39, 40]. Thus, empirical data suggest that current measures used to assess disease activity, such as ACR criteria for remission [41] and DAS28 [42], may not be sensitive enough to exclude lower levels of ongoing disease activity. In future, feasible imaging measures may be part of remission criteria. Based on our results, it may be indicated that the use of the numeric value of CF as a measure of inflammatory activity is relevant and may be a useful for the prediction of the response to anti-TNF-α therapy in RA.

Rheumatology key messages

- Colour US predicts treatment success with anti-TNF-α in patients with RA.
- DAS28 do not predict treatment success with anti-TNF-α in patients with RA.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Kristensen LE, Christensen R, Bliddal H, Geborek P, Danneskiold-Samsoe B, Saxne T. The number needed to treat for adalimumab, etanercept, and infliximab based on ACR50 response in three randomized controlled trials on established rheumatoid arthritis: a systematic literature review. Scand J Rheumatol 2007;36:411–7.

2 Singh JA, Christensen R, Wells GA et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. Can Med Assoc J 2009;181:787–96.

3 Hetland ML, Christensen IJ, Tarp U et al. Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. Arthritis Rheum 2010; 62:22–32.

4 Hyrich KL, Watson KD, Silman AJ, Symmons DP. Predictors of response to anti-TNF-alpha therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Rheumatology 2006;45:1558–65.

5 Brown AK, Quinn MA, Karim Z et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54:3761–73.

6 Terslev L, Torp-Pedersen S, Qvistgaard E, Danneskiold-Samsoe B, Bliddal H. Estimation of inflammation by Doppler ultrasound: quantitative changes after intra-articular treatment in rheumatoid arthritis. Ann Rheum Dis 2003;62:1049–53.

7 Naredo E, Gamero F, Bonilla G, Uson J, Carmona L, Laffon A. Ultrasonographic assessment of inflammatory activity in rheumatoid arthritis: comparison of extended versus reduced joint evaluation. Clin Exp Rheumatol 2005; 23:881–4.

8 Hameed B, Pilcher J, Heron C, Kiely PD. The relation between composite ultrasound measures and the DAS28 score, its components and acute phase markers in adult RA. Rheumatology 2008;47:476–80.

9 Naredo E, Collado P, Cruz A et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis Rheum 2007;57:116–24.

10 Terslev L, Torp-Pedersen S, Savnik A et al. Doppler ultrasound and magnetic resonance imaging of synovial inflammation of the hand in rheumatoid arthritis: a comparative study. Arthritis Rheum 2003;48: 2434–41.

11 Szkudlarek M, Narvestad E, Klarlund M, Court-Payen M, Thomsen HS, Ostergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. Arthritis Rheum 2004;50:2103–12.

12 Walther M, Harms H, Krenn V, Radke S, Faehndrich TP, Gohlke F. Correlation of power Doppler sonography with vascularity of the synovial tissue of the knee joint in patients with osteoarthritis and rheumatoid arthritis. Arthritis Rheum 2001;44:331–8.

13 Ellegaard K, Torp-Pedersen S, Terslev L, Danneskiold-Samsøe B, Henriksen M, Bliddal H. Ultrasound colour Doppler measurements in a single joint as measure of disease activity in patients with rheumatoid arthritis - assessment of concurrent validity. Rheumatology 2009;48:254–7.

14 Taylor PC, Steuer A, Gruber J et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis. Arthritis Rheum 2004;50: 1107–16.

15 Iagnocco A, Perella C, Naredo E et al. Etanercept in the treatment of rheumatoid arthritis: clinical follow-up over one year by ultrasonography. Clin Rheumatol 2008;27: 491–6.

16 Fiocco U, Ferro F, Vezzu M et al. Rheumatoid and psoriatic knee synovitis: clinical, grey scale, and power Doppler ultrasound assessment of the response to etanercept. Ann Rheum Dis 2005;64:899–905.

17 Ribbens C, Andre B, Marcelis S et al. Rheumatoid hand joint synovitis: gray-scale and power Doppler US quantifications following anti-tumor necrosis factor-alpha treatment: pilot study. Radiology 2003;229:562–9.

18 Filippucci E, Iagnocco A, Salaffi F, Cerioni A, Valesini G, Grassi W. Power Doppler sonography monitoring of synovial perfusion at wrist joint in rheumatoid patients treated with adalimumab. Ann Rheum Dis 2006;65: 1433–7.

19 Iagnocco A, Filippucci E, Perella C et al. Clinical and ultrasonographic monitoring of response to adalimumab treatment in rheumatoid arthritis. J Rheumatol 2008;35: 35–40.

20 Terslev L, Torp-Pedersen S, Qvistgaard E et al. Effects of treatment with etanercept (Enbrel, TNRF:Fc) on rheumatoid arthritis evaluated by Doppler ultrasonography. Ann Rheum Dis 2003;62:178–81.

21 Naredo E, Moller I, Cruz A, Carmona L, Garrido J. Power doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. Arthritis Rheum 2008;58: 2248–56.

22 Fleming A, Crown JM, Corbett M. Incidence of joint involvement in early rheumatoid arthritis. Rheumatol Rehabil 1976;15:92–6.

23 Ejbjerg B, McQueen F, Lassere M et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the wrist joint. Ann Rheum Dis 2005; 64(Suppl. 1):i23–47.

24 Johnston MV, Keith RA, Hinderer SR. Measurement standards for interdisciplinary medical rehabilitation. Arch Phys Med Rehabil 1992;73:S3–23.

25 Ostergaard M, Wiell C. Ultrasonography in rheumatoid arthritis: a very promising method still needing more validation. Curr Opin Rheumatol 2004;16:223–30.

26 Arnett FC, Edworthy SM, Bloch DA et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31:315–24.

27 Wells GA, Becker JC, Teng J et al. Validation of the disease activity score 28 (DAS28) and EULAR response criteria based on CRP against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on ESR. Ann Rheum Dis 2009;68: 954–60.

28 Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. Ann Rheum Dis 2008;67:143–9.

29 Ellegaard K, Torp-Pedersen S, Lund H et al. Quantification of colour Doppler activity in the wrist in patients with rheumatoid arthritis - the reliability of different methods for image selection and evaluation. Ultraschall Med 2008;29: 393–8.

30 Qvistgaard E, Rogind H, Torp-Pedersen S et al. Quantitative ultrasonography in rheumatoid arthritis: evaluation of inflammation by Doppler technique. Ann Rheum Dis 2001;60:690–3.

31 Galindo-Rodriguez G, vina-Zubieta JA, Russell AS, Suarez-Almazor ME. Disappointing longterm results with disease modifying antirheumatic drugs. A practice based study. J Rheumatol 1999;26:2337–43.

32 Choy EH, Smith C, Dore CJ, Scott DL. A meta-analysis of the efficacy and toxicity of combining disease-modifying anti-rheumatic drugs in rheumatoid arthritis based on patient withdrawal. Rheumatology 2005;44: 1414–21.

33 Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. Med Care 2001;39:1039–47.

34 Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. J Rheumatol 1998;25:198–9.

35 Aletaha D, Landewe R, Karonitsch T et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Ann Rheum Dis 2008;67:1360–4.

36 Agarwal SK, Glass RJ, Shadick NA et al. Predictors of discontinuation of tumor necrosis factor inhibitors in patients with rheumatoid arthritis. J Rheumatol 2008;35: 1737–44.

37 Terslev L, Torp-Pedersen S, Bang N, Koenig MJ, Nielsen MB, Bliddal H. Doppler ultrasound findings in healthy wrists and finger joints before and after use of two different contrast agents. Ann Rheum Dis 2005;64: 824–7.

38 Ellegaard K, Torp-Pedersen S, Holm CC, Danneskiold-Samsøe B, Bliddal H. Ultrasound in finger joints: findings in normal subjects and pitfalls in the diagnosis of synovial disease. Ultraschall Med 2007;28:401–8.

39 Wakefield RJ, Freeston JE, Hensor EM, Bryer D, Quinn MA, Emery P. Delay in imaging versus clinical response: a rationale for prolonged treatment with anti-tumor necrosis factor medication in early rheumatoid arthritis. Arthritis Rheum 2007;57:1564–7.

40 Ozgocmen S, Ozdemir H, Kiris A, Bozgeyik Z, Ardicoglu O. Clinical evaluation and power Doppler sonography in rheumatoid arthritis: evidence for ongoing synovial inflammation in clinical remission. South Med J 2008;101: 240–5.

41 Felson DT, Anderson JJ, Boers M *et al*. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995; 38:727–35.

42 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.