

RESEARCH ARTICLE

Open Access



Can baseline ultrasound results help to predict failure to achieve DAS28 remission after 1 year of tight control treatment in early RA patients?

D. F. Ten Cate^{1*}, J. W. G. Jacobs², W. A. A. Swen³, J. M. W. Hazes¹, M. H. de Jager⁴, N. M. Basoski⁵, C. J. Haagsma⁶, J. J. Luime^{1†} and A. H. Gerards^{7†}

Abstract

Background: At present, there are no prognostic parameters unequivocally predicting treatment failure in early rheumatoid arthritis (RA) patients. We investigated whether baseline ultrasonography (US) findings of joints, when added to baseline clinical, laboratory, and radiographical data, could improve prediction of failure to achieve Disease Activity Score assessing 28 joints (DAS28) remission (<2.6) at 1 year in newly diagnosed RA patients.

Methods: A multicentre cohort of newly diagnosed RA patients was followed prospectively for 1 year. US of the hands, wrists, and feet was performed at baseline. Clinical, laboratory, and radiographical parameters were recorded. Primary analysis was the prediction by logistic regression of the absence of DAS28 remission 12 months after diagnosis and start of therapy.

Results: Of 194 patients included, 174 were used for the analysis, with complete data available for 159. In a multivariate model with baseline DAS28 (odds ratio (OR) 1.6, 95% confidence interval (CI) 1.2–2.2), the presence of rheumatoid factor (OR 2.3, 95% CI 1.1–5.1), and type of monitoring strategy (OR 0.2, 95% CI 0.05–0.85), the addition of baseline US results for joints (OR 0.96, 95% CI 0.89–1.04) did not significantly improve the prediction of failure to achieve DAS28 remission (likelihood ratio test, 1.04; $p = 0.31$).

Conclusion: In an early RA population, adding baseline ultrasonography of the hands, wrists, and feet to commonly available baseline characteristics did not improve prediction of failure to achieve DAS28 remission at 12 months.

Trial registration: Clinicaltrials.gov, NCT01752309. Registered on 19 December 2012.

Keywords: Ultrasonography, Arthritis, Rheumatoid, Prognosis, Remission, Logistic models, Prediction

Background

The prognosis of patients with rheumatoid arthritis (RA) has improved considerably over the past decades, with early intensive treatment with conventional synthetic and biological disease-modifying anti-rheumatic drugs (DMARDs) [1, 2]. Early intensive treatment nowadays is according to ‘treat to target’ [3] and ‘tight control’

paradigms and strategies [4–6]. This has been shown to beneficially alter the long-term outcome of RA patients, especially if applying these strategies early and within the so-called ‘window of opportunity’ [7–11].

Personalised medicine in a ‘treat to target’ and ‘tight control’ setting in RA would require precise and valid instruments predicting which treatment strategy is necessary to achieve the treatment target. These instruments would enable us to decide which patients need more intensive therapy at the initiation of such therapy. Some patients failing to achieve the treatment target of ‘low-disease activity’ (LDA) or remission would have benefitted from combination therapy from the moment

* Correspondence: d.tencate@erasmusmc.nl

†Equal contributors

¹Department of Rheumatology, Erasmus University Medical Center, Room Na-609; 0031-10-7038251, PO Box 2040, 3000, CA, Rotterdam, The Netherlands

Full list of author information is available at the end of the article



of diagnosis, while monotherapy with methotrexate suffices in others. Not many indisputable prognostic parameters in RA patients exist [12]. The parameters most unequivocally predicting a less favourable outcome of RA are high baseline disease activity [13], rheumatoid factor (RF) positivity [14], and peri-articular bone oedema on magnetic resonance imaging (MRI) [15]. Arthritis assessed at ultrasonography (US) at baseline might also have predictive abilities since US has been shown to predict progression to RA in undifferentiated arthritis [16, 17]. Furthermore, in RA patients in remission, US results predict the occurrence of flare and erosive progression [18–21]. Based on these results, it has been suggested that we should include joint US findings in remission criteria for RA [22, 23], or to use US for classification criteria of RA [24].

A recent study investigating US predictors for clinical remission did not show US to be predictive for clinical remission. However, in this study there was a more than 50% dropout rate, with significant differences between the tender joint count (TJC) and swollen joint count (SJC) of dropouts and patients included in the final analysis [25]. Thus, in a newly diagnosed RA population starting treatment, it remains equivocal whether baseline US results can help to predict clinical outcomes, as assessed by the Disease Activity Score assessing 28 joints (DAS28) [26–28].

The aim of this study was to see whether a more individualised therapy from the start might be possible in newly diagnosed RA patients who will be treated to target; this is achieved by investigating whether failure to achieve DAS28 remission after 1 year is associated with baseline joint US arthritis findings when added to baseline clinical, laboratory, and radiographical predictors. Secondary aims include associations of joint US arthritis findings with the absence of the EULAR good response criterion after 1 year of therapy and with progression of radiographical joint damage. Additionally, we investigated the association between US joint inflammation at baseline and after 1 year.

Methods

This was a multicentre ($n = 7$) Dutch study prospectively following for 1 year a cohort of consecutively recruited, newly diagnosed RA patients (1987 ACR criteria) [29], consisting of four sub-cohorts of which two were ongoing clinical trials: the tREACH (Treatment in the Rotterdam Early Arthritis Cohort), with monitoring every 3 months and treatment escalation if the target of low disease activity was not met [30]; and the U-Act-Early trial [31], with monitoring every 4 weeks and treatment escalation as long as remission was not achieved. The third sub-cohort from which patients were included was the “Reumatologie Online Monitoring Applicatie”

(ROMA), with monitoring intervals varying from 4 to 12 weeks and treatment escalation according to protocol, as long as remission was not achieved [32]. The fourth sub-cohort consisted of early RA patients from the other centres with monitoring at least every 3 months and treatment escalation according to treat to target, at the discretion of the physician, but following the tREACH protocol (cohort one) as closely as possible. Thus, all patients were treated according to a step-up ‘tight control’ and ‘treat to target’ strategy; treating physicians were blinded to results of US examination. Symptom duration was less than 1 year, and all patients were naive for treatment with DMARDs, including biologicals or glucocorticoids. Treatment strategies included at least a potent conventional synthetic DMARD, such as methotrexate and/or a biological DMARD, with or without glucocorticoids. Prednisolone monotherapy was accepted in nine patients. Exclusion criteria were contraindications for conventional synthetic DMARDs, biological therapy, or glucocorticoids, and illiteracy and personality disorders limiting participation in the study. Patients underwent clinical, laboratory, radiographical, and US joint examination at study enrolment (baseline) and at 1-year follow-up. The study was approved of by the medical ethical committee of the Erasmus MC in Rotterdam. All patients signed a consent form prior to inclusion in the study.

Clinical, laboratory, and radiographical assessment

Demographics were recorded for each patient. At baseline and at 1 year joint counts included in DAS [33], as well as the patients’ overall assessment of disease activity using a visual analogue scale (VAS), were scored. Functional ability was evaluated using the Dutch Health Assessment Questionnaire (DHAQ) [34] (https://www.nvr.nl/wp-content/uploads/2014/11/NL_consensus_HAQ.pdf), with higher DHAQ scores indicating more disability.

C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) were obtained at both visits; RF and anti-cyclic citrullinated peptide (aCCP) status were recorded at baseline only. Manufacturer’s cut-off values were used for analyses.

The Sharp-van der Heijde (SvH) scoring method was used to assess radiographic progression of the joints of hands and feet. Anonymised radiographs were read in chronological order [35] by two readers; agreement between them was calculated with intraclass correlation coefficient ICC(A,1). The SvH score consists of the erosion score (0–280) and joint space narrowing (JSN) score (0–168) of hands, wrists, and feet, leading to a maximal score of 448. Progression was defined as a change over time greater than the smallest real difference (SRD), which was derived from the agreement analysis (ICC) between the two readers [36].

US assessment

For the US examination, each centre used a Esaote MyLab60 machine with a LA-435 probe (linear array 6–18 MHz). The second to fifth metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints of both hands were scanned in the dorsal and palmar orientation. Bilaterally, radiocarpal, intercarpal, and second to fifth metatarsophalangeal (MTP) joints were scanned dorsally. The ulnocarpal joint was not included because it was found to be the least reliable orientation in which to scan the wrist [37]. Patient and probe positioning were according to EULAR guidelines [38], with the probe position on the wrist specified in more detail (see Fig. 1). Greyscale US (GSUS) and power Doppler US (PDUS) were performed. The GSUS scoring system for synovitis was semiquantitative (0–3), with ‘0’ indicating no US signs of inflammation and ‘3’ indicating severe inflammation, using Naredo’s modification of the scoring system of Szkudlarek and colleague [39], combining the criteria for joint effusion and synovial proliferation. For erosions, the binary scoring system presented by Wakefield and colleagues [40] was modified to a semi-quantitative (0–3) scoring system, assigning higher scores to larger erosions. PDUS scores were according to the semiquantitative (0–3) PD scoring system of Szkudlarek and colleagues [39].

For PDUS, the colour gain was set at the disappearance of colour noise, frequency at 10 MHz, a pulse repetition frequency (PRF) of 750 Hz and wall filter (WF) at level 3 out of 5. We adjusted the size and position of the colour box to include the subcutaneous tissue to recognize artefacts caused by vessels above the joint.



Fig. 1 Probe position for scanning the wrist. Bottom two red crosses represent Lister’s tubercle (radial) and the ulnar radial joint (ulnar). The top red cross represents the base of the third digit. Green line shows probe position

Based on data in the literature, we considered a joint to be inflamed if it was scored with at least a grade 2 in the GSUS domain and/or at least a grade 1 in the PDUS domain, which resulted in a US joint count (USJC; score 0–28) per patient [41, 42]. For explorative purposes, we investigated different US cut-offs for arthritis in individual joints (e.g. GSUS > 2 and PD > 1) to produce different US sum scores for individual patients.

All seven ultrasonographers, blinded to clinical findings, had been certified for US by the Dutch College of Rheumatology. Prior to the start of the study we organized meetings with all ultrasonographers to discuss the scoring of static images using the scoring systems mentioned above. From these meetings a consensus was reached on how to score images. With this consensus we then composed an atlas with sample images and cartoons for each grade of GSUS, PDUS, and erosive changes. This was distributed to all ultrasonographers, and four training sessions prior to and five more during the study were organized to optimise both interpretation and acquisition reliability between ultrasonographers using this atlas [43].

Statistical analysis

In our protocol the number of patients with a DAS44 > 3.4 1 year after diagnosis had been chosen as the primary outcome. However, actual data showed that only 10 cases met this criterion. Therefore, we redefined our primary outcome by the absence of DAS28 remission (i.e. DAS28 ≥ 2.6) at 1 year of therapy. Clinical, laboratory, radiographical, US sum scores, and US cut-offs were investigated for association with the absence of DAS28 remission (DAS28 ≥ 2.6) at 1 year using a logistic regression of complete cases. Parameters that were univariately associated with a p value < 0.2, or had shown to have prognostic value in the literature, were included in the multivariate model, which was reduced by backward selection ($p < 0.05$). The analyses were repeated after multiple imputation (10 datasets) (results are presented in Additional file 1).

Secondary outcomes were ‘absence of EULAR good response’, analysed with the same analysis as for the primary outcome. Also, ‘presence of radiographical progression’ as assessed by progression on the SvH score greater than the SRD was analysed using a Mann-Whitney U test. Additionally, we investigated the association of the extent of inflammation at US at baseline with that at 1-year follow-up. The cross-sectional associations of US arthritis scores and DAS28 were also evaluated at baseline and 1 year using a Spearman correlation coefficient.

In addition, several US sum scores and US cut-offs were tested in multivariate analyses, while keeping the other variables constant. A value of $p < 0.05$ was

considered to be statistically significant; all tests were two-sided, and STATA 12, (StataCorp, Texas, USA) was used to analyse the data.

Results

Patients

Out of the 174 patients, 159 patients had complete follow-up data (91%); 14 patients dropped out, and for one patient we could not calculate the primary outcome. No significant differences in baseline characteristics were observed between complete ($n = 159$) and incomplete ($n = 15$) cases.

Of the 174 participating patients, 64% were female; the mean (standard deviation (SD)) age was 54.8 (14) years, the median (interquartile range (IQR)) symptom duration was 4.6 (2.8–7.9) months; baseline mean (SD) DAS28 was 4.9 (1.3); 67% of patients were RF positive and 60% were aCCP positive, and 24% of patients had erosions (SvH erosion score > 1) at baseline (Table 1). At 12 months, the mean (SD) DAS28 was 3.5 (0.9) with a median (IQR) of 1 (0–3) swollen joints and 3 (0–5) tender joints; 38% of the patients failed to achieve DAS28 remission and 30% did not achieve the EULAR good response criterion while 3.3% progressed radiographically more than the SRD (6.4 SvH units).

US findings

At baseline ($n = 174$) the median (IQR) joint count per patient regarding synovitis at US (USJC) was 5 (2–9); for erosions at US it was 0 (0–1). The joints most commonly inflamed (GSUS > 1 and/or PDUS > 0) at baseline were radiocarpal joints (right 44%, left 41%), the left

MTP2 (36%), and the left MCP2 on the palmar side (35%). Erosions at US were mainly seen in both MTP5 joints (right 23%, left 22%). On x-ray, erosions were seen in 8% and 15% of these respective joints. At 1 year ($n = 159$), median (IQR) USJC score was 0 (0–3) and for US erosions it was 0 (0–1). The most frequently inflamed joints at 1 year were radiocarpal joints (right 14%, left 15%) and the right MCP2 joint (14%). Erosions were found most frequently in MTP5 joints (right 21%, left 25%). On x-ray, erosions were seen in 15% and 21% of these respective joints. Of the 159 patients for whom 1-year US examination was available, 16 patients (10%) did not have inflammation at US at baseline, and 13 (81%) of those also did not have inflammation at US at 1-year follow-up.

We found no cross-sectional association between US inflammation and DAS28 or the individual DAS28 components at baseline or after 1 year. At baseline, the Spearman correlation coefficient for US inflammation score and DAS28 = 0.26, for SJC28 = 0.22, for TJC28 = 0.04, for ESR = 0.24, and for the VAS = 0.12 (all $p > 0.05$). At 1 year, these were similar.

Reliability of US data

After four training sessions prior to the start of our study, interobserver reliability (ICC(A,1)) at the joint level between the seven ultrasonographers was 0.58 (95% confidence interval (CI) 0.45–0.72) for GSUS, and 0.48 (0.34–0.64) for PDUS, while it was 0.58 (0.49–0.68) for the two modalities combined. However, the percentage exact agreement at the patient level (i.e. the agreement of ultrasonographers for classifying a patient as having arthritis or not) was 95%.

Factors associated with failure of DAS28 remission

Sixty-one patients (38%) failed to achieve DAS28 remission at 12 months. This was associated in the univariate analysis with a higher DAS28 score and higher ESR at baseline ($p < 0.05$). Neither the USJC (GSUS > 1 and/or PDUS > 0) nor US erosion score were associated. Categorizing patients into quartiles based on baseline USJC did not yield significant results, nor did using other definitions of US inflammation. Building a multivariate model resulted in a final model with baseline DAS28, RF, age, and the monitoring strategy with the shortest assessment interval (i.e. every 4 weeks) (Table 2). Adding US results to the final model did not significantly improve the fit of the model (LR $\text{Chi}^2 = 1.04$; $p = 0.31$). The analysis with imputed data yielded almost equal results (see Additional file 1).

Effect of different US sum scores and cut-offs on the multivariate analysis

In our primary analysis, joints were considered to be inflamed at US if GSUS > 1 and/or PDUS > 0 . We also

Table 1 Baseline characteristics of the patients

Age in years, mean (SD)	54.8 (14)
Symptom duration in months ($n = 166$)	4.6
Gender, % female	64%
DAS28, mean (SD)	4.9 (1.3)
SJC28 at baseline	6 (3–11)
TJC28 at baseline	6 (2–10)
ESR ($n = 173$)	27 (12–47)
VAS, mean (SD)	51 (24)
RF, % positive ($n = 171$)	67%
aCCP, % positive ($n = 171$)	60%
RF and aCCP, % positive ($n = 171$)	54%
HAQ ($n = 155$),	0.88 (0.38–1.25)
USJC	5 (2–9)

$n = 174$, unless specified otherwise

Median (interquartile range) are presented unless specified otherwise
aCCP anti-cyclic citrullinated peptide, DAS Disease Activity Score, ESR erythrocyte sedimentation rate, HAQ Health Assessment Questionnaire, RF rheumatoid factor, SD standard deviation, SJC swollen joint count, TJC tender joint count, USJC ultrasonographic joint count, VAS visual analogue scale

Table 2 Univariate and multivariate association of relevant patient characteristics with the outcome: absence of DAS28 remission at 12 months

Variable	Univariate		Multivariate	
	OR (95% CI)	P value	OR (95% CI)	P value
DAS28 per point	1.34 (1.03–1.73)	0.03	1.60 (1.18–2.2)	0.002
Female gender	1.94 (0.96–3.39)	0.06	1.93 (0.91–4)	0.09
RF positive	1.78 (0.87–3.63)	0.11	2.34 (1.07–5.1)	0.03
aCCP positive	1.54 (0.79–3)	0.21		
RF and aCCP positive	1.59 (0.83–3)	0.17		
Age per year	1.01 (0.99–1.04)	0.24		
Monitoring strategy				
Sub-cohort A	Reference	NA	Reference	NA
Sub-cohort B	0.49 (0.13–1.73)	0.27	0.20 (0.05–0.85)	0.03
Sub-cohort C	1.26 (0.47–3.39)	0.64	0.68 (0.20–2.4)	0.55
Sub-cohort D	1.26 (0.26–1.48)	0.28	0.85 (0.30–2.4)	0.76
Symptom duration (months)	1.0 (0.91–1.1)	0.99		
SvH per point	1.04 (0.93–1.15)	0.48		
Smoking yes/no	1.56 (0.73–3.4)	0.25		
HAQ per point	1.16 (0.71–1.9)	0.55		
USJC per point	0.99 (0.92–1.06)	0.77	0.96 (0.89–1.04)	0.31

Absence of remission = 1

Ultrasound joint count (USJC) = number of joints with at least greyscale (GS) > 1 or power Doppler (PD) > 0

Multivariate model including: DAS28 at baseline, RF, monitor strategy, gender, and USJC

Monitor strategies: Sub-cohort A, monitoring at least every 3 months and treatment escalation at the discretion of the physician; sub-cohort B, monitoring every 4 weeks and treatment escalation if no remission; sub-cohort C, monitoring varies between 4 and 12 weeks and treatment escalation if no remission, prespecified in a protocol [24]; sub-cohort D, monitoring every 3 months and treatment escalation if no low-disease activity [23]

aCCP anti-cyclic citrullinated peptide, CI confidence interval, DAS Disease Activity Score, HAQ Health Assessment Questionnaire, NA not applicable, OR odds ratio, RF rheumatoid factor, SvH Sharp-van der Heijde

evaluated several US sum scores and US cut-offs. No US sum score or cut-off was univariately significantly associated with the outcome. The coefficients of the US parameters did not significantly change either when testing all US sum scores and US cut-offs in a multivariate analysis (Table 3).

Secondary outcomes

We evaluated the prediction of not achieving the EULAR good response criterion, radiological progression, and presence of inflammation at US at 12 months as secondary outcomes; 30% of patients did not achieve EULAR good response. Adding US results to the final model did not significantly improve the fit of the model (LR $\text{Chi}^2 = 0.98$; $p = 0.32$). Only five patients showed radiographical progression larger than the SRD. These patients numerically had a higher grade of US joint inflammation at baseline compared to those without radiographical progression: median (IQR) USJC 8 (7–10) versus USJC 5 (2–9), respectively; this difference did not reach statistical significance ($p > 0.05$). Those with a USJC of 0 at 12 months had a median (IQR) USJC of 4 (1–7) at baseline, whereas those with USJC > 0 at follow-up had a median (IQR) USJC of 6 (4–10) at baseline ($p = 0.003$).

Table 3 Results of logistic regression analyses with different US definitions predicting the absence of DAS28 remission at 1 year

Ultrasound score	OR (95% CI)
Per point of:	
GS	0.99 (0.96–1.02)
PD	0.99 (0.95–1.03)
Number of joints with:	
GS > 1	0.67 (0.27–1.71)
PD > 0	0.80 (0.30–2.16)
GS > 1 and/or PD > 0	0.96 (0.89–1.04)
GS > 1 or PD > 0	0.96 (0.42–2.23)
GS > 2	0.56 (0.27–1.20)
GS > 2 or PD > 0	0.57 (0.27–1.21)
GS > 0 and PD > 1	0.82 (0.39–1.7)
GS > 0 and PD > 2	1.12 (0.41–3.01)
GS > 1 and PD > 2	1.12 (0.41–3.01)
GS > 2 and PD > 1	0.71 (0.32–1.56)
GS > 2 and PD > 2	1.26 (0.37–4.24)
GS > 0 or PD > 0	0.40 (0.12–1.33)

Continuous measures: ultrasound joint count = number of joints with greyscale (GS) > 1 and/or power Doppler (PD) > 0; GS sum = sum of GS scores; PD sum = sum of PD scores; GS > x, presence of a joint(s) with a GS > x; PD > y, presence of a joint with a PD > y

CI confidence interval, OR odds ratio

Discussion

Results of routinely performed US at baseline show no additional prognostic information when added to clinical, laboratory, and radiographical parameters in predicting failure to achieve DAS28 remission at 1 year in early RA patients treated to target. The absence of remission was associated with high DAS28 at the time of diagnosis and RF positivity, confirming data in the literature [12–14]. The monitoring strategy with the shortest assessment interval (4-week visits), corrected for DAS28 and RF positivity at baseline, appeared to increase the chance to achieve remission.

The absence of a predictive value of US in this study contrasts with data showing that US predicts patients with undifferentiated arthritis (UA) progressing to RA [16, 17] and RA patients in remission progressing to flare and erosion [18–21]. Our negative study results may be explained by the earlier observation that intensive therapy diminishes the prognostic value of baseline conventional prognostic factors [8]. Our negative results are in line with earlier observations in a smaller population of RA patients not treated to target, 64% of whom received glucocorticoids prior to inclusion in the study [26]. However, in that study ‘time integrated values’ of PDUS correlated with clinical outcomes, which was not found in our study.

It has been suggested that one of the most important steps we have yet to take in US research in RA is to investigate the added value of US and clinical indices in multicentre trials for clinical and radiological outcomes of RA patients [44]. Our study has done this. Although our study has shown that the added predictive value of routinely performed US at baseline in early RA patients is absent if treated to target, US still has a place in personalised medicine for specific indications. For example, in patients for whom the widely used DAS28 is less valid, e.g. patients with concomitant fibromyalgia who will have a high tender joint count and high VAS, and also for obese patients in whom the clinical joint examination may be difficult and who may have a higher ESR without joint inflammation [45]. In these patients, US may really make a clinically relevant contribution. US might help to distinguish patients with or without joint inflammation making more personalised medicine possible. To that aim, new trials should be performed using US measures as target for treatment, or using US as a basis for treatment escalation and tapering; initiatives using US as an outcome measure have been published [46, 47]. In our opinion, the observation that more than 80% of the patients who did not have US inflammation at baseline also did not have US inflammation at follow-up deserves further investigation.

A limitation of our study was heterogeneity in the strategies and assessment intervals of the sub-cohorts from which we included patients. All patients, however, were treated according to a ‘treat to target’ and ‘tight control’ strategy’. Our patients had relative mild disease with a median six swollen joints at baseline; however, this reflects an early RA population and application of the more sensitive, newest RA classification criteria [24]. Both are aspects of normal daily clinical practice in which future US trials as suggested above should also be performed [44]. Another limitation may be the use of US equipment that now might be seen as midrange and less sensitive compared to current high-end machines. However, this also reflects routine practice. More importantly, however, the PD modality in the machines used is actually very sensitive, even better than that of professional research machines [48]. Another possible limitation of our study is that we did not scan all joints included in the DAS28; it could be argued that we missed US inflammation. However, more than 90% of the patients in our study had US inflammation at baseline, and US scoring of a limited sets of joints also well reflects RA activity [49, 50]. A last issue of critical discussion is the moderate agreement coefficients between ultrasonographers, despite organizing several trainings sessions to increase reliability. Generally, reliability in ultrasound is not a bigger problem than in other imaging modalities if performed by trained examiners [51], as was the case in our study. That the agreement coefficients are moderate is probably largely attributable to relatively low levels of detectable inflammation at US in the patients participating in the trainings sessions; ICC values are low in samples with a homogeneous distribution, i.e. with only very high or very low scores of arthritis. Since the heterogeneity of the scores in the patient population largely determines the value of reliability measures [43], we tried to select joints from the patients for the training sessions with scores ranging from 0 to 3. Apparently, we did not succeed in doing this well enough. The hypothesis that the reliability coefficients in our study underestimate reliability is corroborated by the observation that the percentage exact agreement at the patient level was 95%.

Conclusion

In conclusion, in an early RA population, treated according to the ‘treat to target’ paradigm, adding data from US of the joints at baseline to the work-up did not improve prediction for not achieving DAS28 remission or EULAR good response, nor the prediction of radiographical progression at 1 year.

Additional file

Additional file 1: Description of data: univariate and multivariate logistic regression in the multiple imputation dataset (M = 10) using STATA 12 and REALCOM. (DOC 44 kb)

Abbreviations

aCCP: Anti-cyclic citrullinated peptide; DAS: Disease Activity Score; DHAQ: Dutch Health Assessment Questionnaire; DMARD: Disease-modifying anti-rheumatic drug; ESR: Erythrocyte sedimentation rate; GSUS: Greyscale ultrasound; ICC: Intraclass correlation coefficient; IQR: Interquartile range; MCP: Metacarpophalangeal; MTP: Metatarsophalangeal; PDUS: Power Doppler ultrasound; RA: Rheumatoid arthritis; RF: Rheumatoid factor; ROMA: Reumatologie Online Monitoring Applicatie; SD: Standard deviation; SJC: Swollen joint count; SRD: Smallest real difference; SvH: Sharp-van der Heijde; TJC: Tender joint count; tREACH: Treatment in the Rotterdam Early Arthritis Cohort; US: Ultrasonography; USJC: Ultrasound joint count; VAS: Visual analogue scale

Acknowledgements

We thank Dr. Naredo and the Ultrasound School of the Spanish Society of Rheumatology for use of their GSUS scoring system.

Funding

This work was supported by an unrestricted grant from Roche Nederland B.V. Roche had no involvement in the study design, in collection, analysis and interpretation of data, writing of the report, or the decision to submit for publication. The corresponding author had full access to all data and had final responsibility for the decision to submit for publication. The US machines were granted by (in alphabetical order) Abbott, MSD, Pfizer, Schering Plough, and UCB. These companies did not have any involvement in the study design, in collection, analysis and interpretation of data, writing of the report, or the decision to submit for publication.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

DFTC: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. JWJG: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. WAAS: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. JMWH: conception and design, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. MHDj: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. NMB: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. CJH: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. JLL: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version. AHG: conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically. Final approval of the submitted version.

Ethics approval and consent to participate

The study was approved by the medical ethical committee of the Erasmus MC in Rotterdam (MEC2009-333). We received all necessary consents from any patients involved in the study, including the consent to participate in the study.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Rheumatology, Erasmus University Medical Center, Room Na-609; 0031-10-7038251, PO Box 2040, 3000, CA, Rotterdam, The Netherlands. ²Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands. ³Rheumatology, Medical Center Alkmaar, Alkmaar, The Netherlands. ⁴Rheumatology, Albert Schweitzer Hospital, Dordrecht, The Netherlands. ⁵Rheumatology, Maasstad Hospital, Rotterdam, The Netherlands. ⁶Rheumatology, Hospital Group Twente, Almelo, The Netherlands. ⁷Rheumatology, Vlietland Hospital Schiedam, Schiedam, The Netherlands.

Received: 4 August 2017 Accepted: 3 January 2018

Published online: 30 January 2018

References

- Lard LR, Visser H, Speyer I, vander Horst-Bruinsma IE, Zwinderman AH, Breedveld FC, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. *Am J Med.* 2001;111:446–51.
- Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med.* 2000;343:1586–93.
- Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet.* 2004;364:263–9.
- Bakker MF, Jacobs JW, Welsing PM, Vreugdenhil SA, van Booma-Frankfort C, Linn-Rasker SP, et al. Early clinical response to treatment predicts 5-year outcome in RA patients: follow-up results from the CAMERA study. *Ann Rheum Dis.* 2011;70:1099–103.
- Singh JA, Saag KG, Bridges Jr SL, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol.* 2016;68:1–26.
- Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis.* 2014;73:492–509.
- Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J, et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum.* 2002;46:894–8.
- Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol.* 2003;21(5 Suppl 31):S154–7.
- Breedveld F. The value of early intervention in RA—a window of opportunity. *Clin Rheumatol.* 2011;30 Suppl 1:S33–9.
- van Nies JA, Krabben A, Schoones JW, Huizinga TW, Kloppenburg M, van der Helm-van Mil AH. What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review. *Ann Rheum Dis.* 2014;73:861–70.
- van Nies JA, Tsonaka R, Gaujoux-Viala C, Fautrel B, van der Helm-van Mil AH. Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? Results on the Leiden early arthritis clinic and ESPOIR cohorts. *Ann Rheum Dis.* 2015;74:806–12.
- Drouin J, Haraoui B, 3e Initiative Group. Predictors of clinical response and radiographic progression in patients with rheumatoid arthritis treated with methotrexate monotherapy. *J Rheumatol.* 2010;37:1405–10.
- Romao VC, Canhao H, Fonseca JE. Old drugs, old problems: where do we stand in prediction of rheumatoid arthritis responsiveness to methotrexate and other synthetic DMARDs? *BMC Med.* 2013;11:17.
- Aletaha D, Alasti F, Smolen JS. Rheumatoid factor determines structural progression of rheumatoid arthritis dependent and independent of disease activity. *Ann Rheum Dis.* 2013;72:875–80.
- Hetland ML, Ejbjerg B, Horslev-Petersen K, Jacobsen S, Vestergaard A, Jurik AG, et al. MRI bone oedema is the strongest predictor of subsequent radiographic progression in early rheumatoid arthritis. Results from a 2-year randomised controlled trial (CIMESTRA). *Ann Rheum Dis.* 2009;68:384–90.

16. Filer A, De Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, et al. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Ann Rheum Dis*. 2011;70:500–7.
17. Salaffi F, Ciapetti A, Gasparini S, Carotti M, Filippucci E, Grassi W. A clinical prediction rule combining routine assessment and power Doppler ultrasonography for predicting progression to rheumatoid arthritis from early-onset undifferentiated arthritis. *Clin Exp Rheumatol*. 2010;28:686–94.
18. Scire CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. *Rheumatology*. 2009;48:1092–7.
19. Peluso G, Michelutti A, Bosello S, Gremese E, Toluoso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Ann Rheum Dis*. 2011;70:172–5.
20. Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, et al. Power Doppler ultrasound, but not low-field magnetic resonance imaging, predicts relapse and radiographic disease progression in rheumatoid arthritis patients with low levels of disease activity. *Arthritis Rheum*. 2012;64:67–76.
21. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum*. 2008;58:2958–67.
22. Haavardsholm EA, Lie E, Lillegraven S. Should modern imaging be part of remission criteria in rheumatoid arthritis? *Best Pract Res Clin Rheumatol*. 2012;26:767–85.
23. Ostergaard M, Moller-Bisgaard S. Rheumatoid arthritis: is imaging needed to define remission in rheumatoid arthritis? *Nat Rev Rheumatol*. 2014;10:326–8.
24. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham 3rd CO, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis*. 2010;69:1580–8.
25. Horton SC, Tan AL, Freeston JE, Wakefield RJ, Buch MH, Emery P. Discordance between the predictors of clinical and imaging remission in patients with early rheumatoid arthritis in clinical practice: implications for the use of ultrasound within a treatment-to-target strategy. *Rheumatology (Oxford)*. 2016;55:1177–87.
26. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, 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.
27. Ellegaard K, Terslev L, Christensen R, Szkudlarek M, Schmidt WA, Jensen PS, et al. Comparison of discrimination and prognostic value of two US Doppler scoring systems in rheumatoid arthritis patients: a prospective cohort study. *Clin Exp Rheumatol*. 2014;32:495–500.
28. Dougados M, Devauchelle-Pensec V, Ferlet JF, Jousse-Joulin S, D'Agostino MA, Backhaus M, et al. The ability of synovitis to predict structural damage in rheumatoid arthritis: a comparative study between clinical examination and ultrasound. *Ann Rheum Dis*. 2013;72:665–71.
29. 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.
30. de Jong PH, Hazes JM, Han HK, Huisman M, van Zeben D, van der Lubbe PA, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy: 1-year data of the tREACH trial. *Ann Rheum Dis*. 2014;73:1331–9.
31. Bijlsma JW, Welsing PM, Woodworth TG, Middelink LM, Petho-Schramm A, Bernasconi C, et al. Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. *Lancet*. 2016;388:343–55.
32. Vermeer M, Kuper HH, Bernelot Moens HJ, Hoekstra M, Posthumus MD, van Riel PL, et al. Adherence to a treat-to-target strategy in early rheumatoid arthritis: results of the DREAM remission induction cohort. *Arthritis Res Ther*. 2012;14:R254.
33. van der Heijde DM, van 't Hof M, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologist. *J Rheumatol*. 1993;20:579–81.
34. Siegert CE, Vleming LJ, Vandenbroucke JP, Cats A. Measurement of disability in Dutch rheumatoid arthritis patients. *Clin Rheumatol*. 1984;3:305–9.
35. van Tuyl LH, van der Heijde D, Knol DL, Boers M. Chronological reading of radiographs in rheumatoid arthritis increases efficiency and does not lead to bias. *Ann Rheum Dis*. 2014;73:391–5.
36. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res*. 2001;10:571–8.
37. Dopazo Gonzalez N, Ten Cate DF, Swen WA, Mera Varela A, Insua Vilarino SA, Perez-Pampin E, et al. The most reliable probe position in the ultrasonographic examination of the wrist in rheumatoid arthritis. *Clin Exp Rheumatol*. 2012;30:566–9.
38. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al. Guidelines for musculoskeletal ultrasound in rheumatology. *Ann Rheum Dis*. 2001;60:641–9.
39. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum*. 2003;48:955–62.
40. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol*. 2005;32:2485–7.
41. Witt M, Mueller F, Nigg A, Reindl C, Leipe J, Proft F, 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.
42. Ten Cate DF, Luime JJ, Hazes JM, Kleinsink GJ, Jacobs JW. Is the frequent sonographic anechoic area distally in metacarpophalangeal joints a sign of arthritis? *Ultrasound Med Biol*. 2014;40:2537–41.
43. Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewe R. Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung et al. *Arthritis Care Res*. 2010;62:1357–8. author reply 1358.
44. Mandl P, Kurucz R, Niedermayer D, Balint PV, Smolen JS. Contributions of ultrasound beyond clinical data in assessing inflammatory disease activity in rheumatoid arthritis: current insights and future prospects. *Rheumatology (Oxford)*. 2014;53(12):2136–42.
45. Jacobs JW, Ten Cate DF, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice. *Rheumatology (Oxford)*. 2015; 54(6):959–61.
46. Dale J, Purves D, McConnachie A, McInnes I, Porter D. Tightening up? Impact of musculoskeletal ultrasound disease activity assessment on early rheumatoid arthritis patients treated using a treat to target strategy. *Arthritis Care Res (Hoboken)*. 2014;66:19–26.
47. Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Ann Rheum Dis*. 2012;71:799–803.
48. Ten Cate DF, Luime JJ, van der Ven M, Hazes JM, Kooiman K, de Jong N, et al. Very different performance of the power Doppler modalities of several ultrasound machines ascertained by a microvessel flow phantom. *Arthritis Res Ther*. 2013;15:R162.
49. Zhou L, Wang G, Liu X, Song J, Chen L, Xu H. Matrix metalloproteinase-3 and the 7-joint ultrasound score in the assessment of disease activity and therapeutic efficacy in patients with moderate to severe rheumatoid arthritis. *Arthritis Res Ther*. 2017;19:250.
50. Naredo E, Rodriguez M, Campos C, Rodriguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum*. 2008;59:515–22.
51. Thiele RG. Ultrasonography applications in diagnosis and management of early rheumatoid arthritis. *Rheum Dis Clin North Am*. 2012;38:259–75.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

