Quantitative power Doppler ultrasound measures of peripheral joint synovitis in poor prognosis early rheumatoid arthritis predict radiographic progression

Dee Sreerangaiah¹,a, Michael Grayer², Benjamin A. Fisher¹,b, Meilien Ho³, Sonya Abraham¹,c and Peter C. Taylor⁴

Abstract

Objective. To assess the value of quantitative vascular imaging by power Doppler US (PDUS) as a tool that can be used to stratify patient risk of joint damage in early seropositive RA while still biologic naive but on synthetic DMARD treatment.

Methods. Eighty-five patients with seropositive RA of <3 years duration had clinical, laboratory and imaging assessments at 0 and 12 months. Imaging assessments consisted of radiographs of the hands and feet, two-dimensional (2D) high-frequency and PDUS imaging of 10 MCP joints that were scored for erosions and vascularity and three-dimensional (3D) PDUS of MCP joints and wrists that were scored for vascularity.

Results. Severe deterioration on radiographs and ultrasonography was seen in 45 and 28% of patients, respectively. The 3D power Doppler volume and 2D vascularity scores were the most useful US predictors of deterioration. These variables were modelled in two equations that estimate structural damage over 12 months. The equations had a sensitivity of 63.2% and specificity of 80.9% for predicting radiographic structural damage and a sensitivity of 54.2% and specificity of 96.7% for predicting structural damage on ultrasonography.

Conclusion. In seropositive early RA, quantitative vascular imaging by PDUS has clinical utility in predicting which patients will derive benefit from early use of biologic therapy.

Key words: early rheumatoid arthritis, treatment, ultrasonography

Introduction

RA is a syndrome characterized by synovial inflammation of variable extent and severity as well as rate of radiographic progression. The most effective inhibition of structural damage is associated with biologic TNF inhibition, but this is also costly. In biologic-naive patients it is difficult to predict who will most benefit from biologic disease modification as judged by disease activity scores alone. In many trials, probability plots of individual radiographic progression demonstrate that most patients receiving MTX alone do not progress [1, 2]. Conversely, some patients meeting 28-joint DAS (DAS28) remission targets...
have subclinical synovitis on imaging and continue to accrue damage [3, 4].

Synovitis in the active phase of erosive disease is metabolically active and vascular. Power Doppler US (PDUS) enables detection of low-velocity blood flow in inflamed small joints of the hands [4, 5] and shows significant correlation to endothelial cell density on synovial histology, permitting reliable assessment of joint inflammation [6, 7] with superior sensitivity to clinical examination [5, 8]. Greyscale sonography may also detect RA erosions earlier than conventional radiography [9, 10]. While three-dimensional (3D) PDUS is a novel technique within the rheumatologic field, where it is yet to be validated, it has been well researched and utilised within the field of obstetrics and gynaecology and has proved to be highly reproducible between observers [11]. Using this technique, we can now assess a virtually reconstructed vascular tree within a volume of interest and can determine its vascularization by calculating indices using VOCAL software (GE Medical Systems, Waukesha, WI, USA). These indices comprise the vascularization index, flow index and vascularization flow index and are thought to reflect the number of vessels within the vascularization index volume of interest, the intensity of flow at the time of the 3D sweep (flow index) and both blood flow and the vascularization flow index. In vivo studies have shown that the vascularization index correlates positively with microvessel density count as assessed by immunohistochemical techniques [12]. In this article we describe the potential of quantitative vascular imaging by PDUS as a tool to stratify the risk of joint damage in early seropositive, biologic-naive RA.

Patients and methods

Eighty-five patients meeting 1987 ACR classification criteria for RA of <3 years duration and seropositive for anti-CCP antibody and/or IgM RF were prospectively recruited and treated according to the standard of care with conventional DMARDs. At each study visit the DAS28, HAQ and Belza fatigue score were recorded along with ESR and CRP. The study was conducted in compliance with the Declaration of Helsinki and ethical approval was obtained from the Hammersmith Research Ethics Committee (08/H0707/114). All subjects gave written informed consent.

Imaging assessments

All patients underwent 2D greyscale US, PDUS of 10 MCP joints over the dorsal surface in the transverse and longitudinal planes and 3D PDUS of 10 MCP and bilateral wrist joints at 0 and 12 months. A single ultrasonographer (D.S.) carried out all scans with a GE Logiq 9 scanner [multifrequency linear (7-12 MHz)] and 4D16L3D probes.

All images were anonymized and stored for subsequent analysis using a computerized image analysis system. The 2D images were scored separately for synovial thickening and vascularity against an analogue scale from 0 to 4 (0, no hypertrophy or vascularity; 1, minimal; 2, mild; 3, moderate; 4, severe hypertrophy or vascularity).

Both hypertrophy and vascularity on 2D images were also calculated by pixel count in a defined region of interest for each joint. Finally, the number of erosions in each MCP joint was determined on greyscale images, with an erosion defined as an intra-articular discontinuity of the bone surface that is visible in two perpendicular planes. The 3D scans were performed using the GE 4D16L3D volumetric probe using an automated sweep and images were scored semi-quantitatively for vascularity and also quantitatively in cubic millimetres using LOGIQworks software (GE Healthcare, Waukesha, WI, USA).

Statistical methodology

Data were obtained for 85 patients for a total of 25 variables categorized into five categories: US scans [13], where each US variable was the sum score across all 10 MCP joints; physiological risk factors for RA [6]; demographic variables [2]; timing/baseline variables [3] and outcome variables (2). Outcomes were radiographic progression and US progression. For radiographic progression we defined severe deterioration as a change in the van de Heijde–Sharp (vdHS) score of 4 or more since this would equate with complete ankylosis or luxation of a joint. For US progression we defined a change in US erosion score of two or more as denoting severe deterioration, as these were only scored at MCP joints. Values that were right-skewed were transformed so that extreme data points did not exert undue influence on the regression analyses. Logarithmic transformation was preferred as standard. For variables with many zero values, a square root transformation was used to avoid loss of data.

For each of two outcome variables, analyses were conducted in three parts. First, the bivariate relationship between the outcome variable and each US scan variable was evaluated using linear regression. The 2D and 3D US variables with the best correlation with the outcome variable were selected to be carried forward to the second part (compared using $R^2$). Second, an exploratory multivariate analyses was conducted using the two US variables selected from part 1 and each of the other variables.

Finally, a plausible multivariate model was arrived at by starting with a linear regression model including all variables carried forward from part 2 and eliminating variables one by one based on the highest $P$-value. A best model was arrived at through this iterative process using the adjusted $R^2$ score as a diagnostic criterion (using the formula adjusted $R^2 = 1 - (1 - R^2)\frac{(N - 1)}{(N - k - 1)}$), where $N$ is the number of observations and $k$ is the number of predictor variables).

The validity and prognostic utility of each model were assessed by evaluating its sensitivity and specificity in predicting deterioration in the outcome variable. All analysis was conducted using Stata 11 for Windows (StataCorp, College Station, TX, USA).

Results

Of the 85 patients [median age 50 years (range 18–79), median disease duration 12 months], 64 (75%) were...
female. The 1 year median change in vdHS score was 2 (interquartile range [IQR] 0–8]. Severe radiographic deterioration was seen in 25 (39%) female and 12 (60%) male patients. Severe US deterioration was seen in 16 (25%) female and 8 (38%) male patients.

**Radiographic progression**

The 3D and 2D US parameters with the strongest bivariate association with change in vdHS score were 3D power Doppler volume ($R^2 = 0.34$; illustrated in Fig. 1) and 2D longitudinal vascularity score ($R^2 = 0.33$). Four additional baseline variables were identified in the exploratory multivariate analysis for potential inclusion in the full multivariate model: baseline vdHS score, CRP, ESR and gender. Following the full multivariate analysis, CRP and the 2D longitudinal vascularity score were eliminated from the final model. Although the association of gender and change in vdHS score did not meet usual bounds for statistical significance ($P = 0.106$), its exclusion reduced model fit. We decided to evaluate the sensitivity and specificity of two models with gender included and excluded (Table 1).

**US progression**

The 3D US parameter with the strongest bivariate association with change in US erosion score was 3D power Doppler volume ($R^2 = 0.38$). The 2D transverse vascularity score ($R^2 = 0.32$) was the strongest correlate of the vascularity indices; the correlation of the other 2D US scans with erosion progression was not considered to be sufficiently greater to outweigh the practical advantages of the simpler vascularity index.

Four additional baseline variables were identified in the exploratory multivariate analysis for potential inclusion in the full multivariate model: age, baseline US erosion score, gender and CRP. Following the full multivariate analysis, CRP and the 2D transverse vascularity score were eliminated from the final model, for which sensitivity and specificity were evaluated. For both radiographic and US progression, the high specificity of the equations indicates that the test has utility for stratification of patients who will experience severe structural progression over 1 year.

**Reliability analysis**

The method used was the intraclass correlation coefficient (ICC). For 2D measures, 30 US images were chosen at random and rescored blindly at the end of the study by the original scanner. Of the eight 2D parameters, two scored poorly with an ICC <0.5 (these were longitudinal synovial thickness measured semi-quantitatively and transverse power Doppler area measured quantitatively). The remaining six parameters showed excellent reliability, with ICCs between 0.79 and 0.96, with four of the six parameters having an ICC >0.9.

For 3D measures, 30 images were again chosen at random and rescored blindly at the end of the study by the original scanner. In addition, a second observer scored those same images independently to evaluate interobserver reliability. The ICC for intraobserver reliability was 1.0 (95% CI 1.0, 1.0) and for interobserver reliability was 0.99 (95% CI 0.99, 1.0).

---

![Fig. 1 Scatterplot of change in vdHS score vs three-dimensional power Doppler volume](https://academic.oup.com/rheumatology/article-abstract/55/1/89/1794889)
**Table 1** Model equations, sensitivity and specificity of selected final models

<table>
<thead>
<tr>
<th>Model number</th>
<th>Equation</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
</tr>
</thead>
</table>
| 1            | \[
\text{Radiographic progression (change in vdHS score)} = -1.61 + 0.433^*\log(\text{power Doppler volume}) + 0.220^\sqrt{\text{vdHS score}} + 0.442^*\log(\text{ESR}) + 0.503^*(\text{gender})
\] | 63.2 (95% CI 47.1, 79.2) | 80.9 (95% CI 69.2, 92.5) |
| 2            | \[
\text{US progression (change in US erosions score)} = -0.892 + 0.419^*\log(\text{power Doppler volume}) + 0.236^\sqrt{\text{baseline vdHS score}} + 0.402^*\log(\text{ESR})
\] | 67.6 (95% CI 51.7, 83.4) | 78.7 (95% CI 66.6, 90.9) |
| 3            | \[
\text{Power Doppler volume} = -1.09 + 0.289^*\log(\text{power Doppler volume}) + 0.0180*(\text{age}) + 0.173^*\sqrt{\text{baseline erosion score}} + 0.314^*(\text{gender})
\] | 54.2 (95% CI 32.7, 75.7) | 96.7 (95% CI 92.1, 100.0) |

*Estimated by the formula as >100%; rounded down to 100%. Power Doppler volume measured in cubic millimetres. Gender coded as 1 (female) or 2 (male). Age measured in years at the time of baseline visit. √: square root; *: multiplied; vdHS: van de Heijde–Sharp.

**Discussion**

We have previously described relationships between power Doppler US and erosive progression in a small early RA cohort [13]. Here we describe two exploratory equations that estimate structural damage in biologic-naive, early RA patients over 12 months, using easily recorded baseline variables. The uniqueness of our models is in the inclusion of baseline US data, in particular the volume of vascularised synovium by PDUS of the MCP joints, which has not been described previously. Vastesaeger et al. [14] developed a risk matrix model for the prediction of RA radiographic progression, but the potential contribution from imaging was not assessed. In a more recent longitudinal study, greyscale and PDUS of the metacarpal, wrist and metatarsal joints was found to provide the optimum US data to improve on clinical predictive models for the development of RA from early undifferentiated inflammatory arthritis [15]. In this study, power Doppler variables performed better than greyscale variables and had a uniquely high specificity for RA compared with other disease classifications. Similarly, other investigators reported that PDUS had good sensitivity and specificity for persistence of inflammatory arthritis in a cohort with early inflammatory symptoms [16] and that ultrasonographic greyscale inflammation is an independent predictor of 1 year MRI erosive progression in early RA, with an odds ratio of 2.01 [17]. A major limitation of this latter study, however, was the lack of PDUS assessment.

Thus our data are in line with previous findings of ultrasonographic power Doppler signal as an independent predictor of structural damage in early RA, which we have defined by two different imaging modalities—conventional radiograph and US. Structural damage on US as seen in our study is likely to be an underrepresentation of true damage because of the technical limitations of erosion detection by US. There are limitations to our study, including that we did not examine the volar surface of the MCP joints for erosions; however, this is less likely to have an impact on the results, as it has been estimated that only 2% of erosions occur at this aspect [18], nor did we examine the wrists or MTP joints for erosive change or vascular signal, which have been shown to be involved in the majority of patients with RA. Despite this, at baseline we still detected twice as many erosions at the MCP joints on US compared with radiographs, which increased to approximately four times as many by 12 months with a significant correlation to baseline vascular signal.

Our analysis suggests that 3D power Doppler volume predicts radiographic progression in RA and has the potential to be used, in conjunction with demographic data, to inform decisions on whether anti-TNF treatment should be deployed immediately. Ours is the first report of biomarker utility of 3D PDUS volume in a clinical setting. We have combined these data in two equations with high specificity for predicting which patients will experience severe deterioration over a year while on conventional DMARDS and would thus benefit from early use of biologic therapy. While vascular signal seen on US has previously been described to predict which patients would remain on anti-TNF therapy after 1 year [19], this is the first time that it has found clinical utility in predicting which patients would derive clinical benefit from the early use of biologic therapy.

**Acknowledgements**

The authors would like to thank Amy Fox and Catherine McClinton for nursing support throughout this study. This study was sponsored by Imperial College London. Peter Taylor thanks Arthritis Research UK for their funding of the Arthritis Research UK Early Arthritis Treatment Centre at the University of Oxford and the National Institute of Health Research (NIHR) for their funding of the NIHR Biomedical Research Centre in Musculoskeletal Disease at Oxford University Hospitals National Health Service (NHS) Trust and the University of Oxford. Sonya Abraham thanks the NIHR for funding the research supported by the NIHR Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the
authors and not necessarily those of Arthritis Research UK, the NHS, the NIHR or the Department of Health.

**Funding:** This work was supported by the Medical Research Council (Imperial College Grant P11339) and AstraZeneca (Imperial College Grant P32116).

**Disclosure statement:** P.C.T. received a grant to the NIHR from AstraZeneca and has also served as a consultant to AstraZeneca. M.H. is an employee of AstraZeneca and holds stock and stock options. M.G. has left Dianthus to work for Phastar Ltd., who provide statistical services to a number of clients, including AstraZeneca. They have contributed to a number of AstraZeneca research and development projects in this capacity. All other authors have declared no conflicts of interest.

**References**


