RHEUMATOLOGY

Rheumatology 2015;54:1897–1903
doi:10.1093/rheumatology/kev205
Advance Access publication 11 June 2015

Original article

# Grading of ultrasound Doppler signals in synovitis: does it need an update?

Wolfgang A. Schmidt[1], Verena Schönau[1,2], Britta E. Reiche[3], Philipp D. Oberdorfer[3], Sarah Ohrndorf[3] and Marina Backhaus[3]

## Abstract

**Objectives.** To compare subjective estimation with computerized quantification of synovial perfusion in active RA, develop new quantitative scores, establish quantitative limit values for the respective grades in order to achieve even distribution and compare the new scores with the established semi-quantitative score.

**Methods.** Patients fulfilling the 2010 RA classification criteria in whom US showed power Doppler signals in one or more wrist or MCP joints were included. Right and left wrists and MCP joints 1–5 were examined with dorsal and volar scans. The proportion of the synovium covered by Doppler signals was estimated and quantified electronically in the area with the greatest fraction of colour signals.

**Results.** Forty-one RA patients [29 females, mean age 62 years (s.d. 14), disease duration 11 years (s.d. 13), 28-joint DAS 5.5 (s.d. 1.3)] were examined. Colour signals were found in 192 of 984 joint regions. Forty-two, 139 and 11 regions were allocated to the semi-quantitative grades 1, 2 and 3, respectively, with electronically calculated colour fractions of 3.9%, 12.6% and 29.7%. The mean estimated colour fractions were lower than the mean measured fractions. An even distribution of the scores was found for estimated colour fractions of >0–10% for grade 1, >10–25% for grade 2 and >25% for grade 3 and for measured colour fractions of >0–6% for grade 1, >6–12% for grade 2 and >12% for grade 3.

**Conclusion.** This study suggests replacing the semi-quantitative grading system for synovial Doppler US with more evenly distributed quantitative scores that might better reflect treatment response.

**Key words:** ultrasound, Doppler ultrasound, rheumatoid arthritis, disease activity, synovitis, hand, grading system, outcome measure.



CLINICAL SCIENCE

**Rheumatology key messages**

- The current synovitis Doppler score has an uneven distribution of grades in active RA.
- Sonographers estimated higher colour fractions for synovial perfusion in RA than the computerized pixel count software.
- More evenly distributed quantitative synovial perfusion scores potentially increase the sensitivity to change for follow-up RA trials.

## Introduction

Doppler US allows detection of the perfusion of synovial tissue in inflammatory rheumatic diseases. It has become an important tool for determining the inflammatory activity in RA and other rheumatic diseases. The presence of Doppler signals represents a prognostic criterion for developing erosive disease in RA [1–3].

Doppler US depends on the Doppler effect. This is a change in the frequency of a wave resulting from the motion of red blood cells. Two different modes of Doppler US are currently used for determining synovial vascular flow: colour Doppler and power Doppler. Power Doppler depicts the amplitude, or power, of Doppler signals rather than the frequency shift. In many modern US machines the sensitivity for detecting Doppler signals is

[1]Immanuel Krankenhaus Berlin, Medical Centre for Rheumatology Berlin-Buch, Berlin, [2]Department of Internal Medicine 3, University of Erlangen-Nuremberg, Erlangen and [3]Department of Rheumatology and Clinical Immunology, Charité–University Medicine Berlin, Berlin, Germany

Submitted 21 October 2014; revised version accepted 23 April 2015

Correspondence to: Wolfgang A. Schmidt, Immanuel Krankenhaus Berlin, Medical Centre for Rheumatology Berlin-Buch, Lindenberger Weg 19, 13125 Berlin, Germany. E-mail: w.schmidt@immanuel.de

similar for the colour Doppler mode compared with the power Doppler mode [4].

The first studies investigating Doppler US have applied a qualitative scoring system for grading intra-articular Doppler signals: 0, no colour signals; 1, mild hyperaemia; 2, moderate hyperaemia; 3, marked hyperaemia [5–8]. Other investigators suggested measuring the resistance index [9] or calculating time–intensity curves after having injected US contrast agents [10]. These methods have not gained acceptance because of their lack of feasibility in daily clinical practice. Otherwise, colour pixels could be counted with external software [1, 9, 11]. Such software is now provided as internal software and should become more widely available in the future.

In 2003 Szkudlarek *et al*. [12] suggested a new semi-quantitative grading system (Table 1). In addition, other authors have suggested limiting the number of visible single vessel signals representing grade 1 to a maximum of three colour spots [13].

Since then this score has been used in the majority of Doppler US studies [14–16]. Studies analysing the inter- and intrarater variability of the above mentioned semi-quantitative Doppler US score have found good agreements. Agreements tend to be higher for the Doppler US score than for grey-scale US findings [17–22]. Effective treatment leads to a reduction in colour signals. The score correlates with the longitudinal assessment of the clinical activity in RA [18, 23–33]. Other trials are currently investigating whether patients in clinical remission might benefit from treatment escalation if US reveals synovial hypervascularity [34, 35].

With advancing US technology, the sensitivity for detecting Doppler signals has considerably increased in recent years. Studies have shown that the majority of inflamed joints are classified as grade 2 according to the Szkudlarek score, while grades 1 and 3 are rare [36, 37]. Grade 2 includes for instance both confluent signals that cover 5% or slightly <50% of the synovium. These findings also reflect the experience of the authors in clinical practice. This uneven distribution between grades may impair the sensitivity to change in follow-up studies using Doppler as an outcome measure. Even with new technology, power Doppler signals rarely cover >50% of the synovium (grade 3). This study aims to test these hypotheses.

**Table 1** Semi-quantitative grading system for Doppler US signals in synovial tissue [12]

| Grade | |
|---|---|
| **0** | No flow in the synovium |
| 1 | Single vessel signals (up to 3) |
| 2 | Confluent vessel signals in less than half of the area of the synovium |
| 3 | Vessel signals in more than half of the area of the synovium |

Adapted from Szkudlarek *et al.* [12] and Naredo *et al.* [13].

US technology that allows Doppler quantification, that is, determining the ratio of the synovial area covered by Doppler signals in relation to the total synovial area, is now available. This technology may facilitate a quantitative score.

The objectives of this study are to compare subjective estimation and computerized quantification of synovial perfusion in active RA, develop a new quantitative score, establish quantitative limit values for the respective grades in order to achieve even distribution and compare the new score with the established semi-quantitative score (0–3) by Szkudlarek *et al*. [12] and modified by Naredo *et al*. [13].

## Patients and methods

Consecutive patients diagnosed with RA fulfilling the 2010 ACR/EULAR RA classification criteria were included in the study [38]. Clinical examination had to reveal at least one swollen MCP or wrist joint. In addition, US had to show synovial hypertrophy with the presence of power Doppler signals in at least one MCP or wrist joint. Patients gave written informed consent according to the Declaration of Helsinki before entering the study. Approval for the study was given by the ethical committee of the Charité University Medicine Berlin, Germany (registration number EA1/069/12).

Forty-one patients were examined in two centres, 30 patients in the Medical Centre for Rheumatology in Berlin-Buch and 11 patients in the Department of Rheumatology of the Charité–University Medicine Berlin. Both wrists and all 10 MCP joints were examined by a physician experienced in musculoskeletal US with dorsal and volar longitudinal scans. These scans included probe positions of up to 45° medially and laterally from the standard dorsal or volar probe position. Thus 24 joint regions of 12 joints were examined in every patient.

All examinations were done with the same GE Logiq E9 XDclear US equipment (GE Healthcare, Wauwatosa, WI, USA) constructed in 2013 and equipped with an 8- to 18-MHz hockey stick probe (L8-18i-D). US machine settings were standardized as follows: (i) grey-scale settings: frequency 18 MHz; frame average 2; line density 2; dynamic range 78; speckle reduction imaging 3; (ii) power Doppler imaging settings: pulse repetition frequency 0.5 kHz; power Doppler frequency 10 MHz; threshold 100%; wall filter 0/32; spatial filter 1; frame average 4; package size 14; line density 2. The power Doppler gain was adjusted just below the appearance of artefacts. The power Doppler box included all visible parts of the joint within the synovial capsule. Settings were optimized to the highest possible colour sensitivity in two separate sessions with an application specialist of the US company before the study. Measurement was done in the probe position that revealed the greatest fraction of power Doppler signals.

Findings were first graded semi-quantitatively (0–3) according to the Szkudlarek score modified to Naredo. The proportion of the synovial area covered by Doppler signals was then subjectively estimated by the sonographer and finally determined with the Doppler quantification software (Q-Analysis, GE Healthcare). The sonographer placed the probe over the area where he or she subjectively saw more colour than in other scan orientations. Then a 3-s video sequence was recorded. Thereafter the sonographer drew a line to indicate the area of the synovium. Subsequently the software presented a graph showing the colour fraction over these 3 s (Fig. 1). The US image corresponding to the highest colour fraction was reviewed in order to exclude artefacts. In the case of flash artefacts, the second or third peak was evaluated if this was free of artefacts.

The sonographers were blinded to further clinical data. However, the clinical phenotype (e.g. joint swelling) was obvious to the sonographers while performing the US examination. Examinations were repeated by a second sonographer blinded to the results of the first sonographer in 15 joint regions of six patients in order to estimate interrater reliability.

For the statistical analysis, all data were recorded in an Excel (Microsoft, Redmond, WA, USA) spreadsheet and analysed by descriptive analysis methods. The numeral and percentage frequencies of the estimation and computerized quantification as well as of the synovial perfusion evaluated by the conventional semi-quantitative grading system were assessed for each joint region and compared with each other. Weighted $\kappa$ was determined to describe interrater reliability for the grading system results in joints that exhibited colour signals (grades 1–3). For the percentage measurements, Lin's concordance correlation coefficient was applied.

## Results

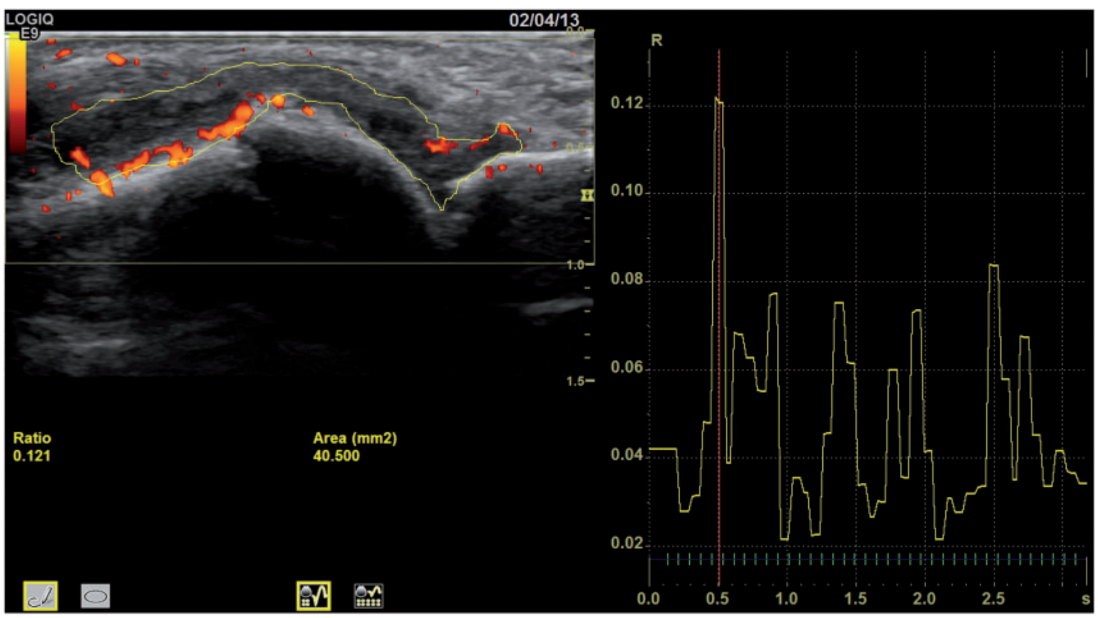
Of the 41 patients, 29 were female. The mean age of all patients was 62 years (S.D. 14). The mean disease duration was 11 years (S.D. 13.9). The mean 28-joint DAS with CRP (DAS28-CRP) was 5.5 (S.D. 1.3). The mean CRP was 25 mg/l (S.D. 40; normal $<$5 mg/l). Radiography of the hands and forefeet revealed erosions in 28 patients. RF was positive in 66% of patients and ACPA was positive in 63%.

Colour signals were found in 192 of 984 joint regions. Doppler signals were most commonly detected in the wrist, followed by the MCP 2 and 3 joints. They were more commonly detectable in dorsal rather than volar scans (Table 2). Table 3 shows the mean subjectively estimated and measured colour fractions related to the joint regions.

The highest subjectively estimated colour fraction was 70% in dorsal scans of a left dorsal wrist joint, a right dorsal MCP 5 joint and a left dorsal MCP 3 joint. The highest measured colour fraction was 52% in the above-mentioned dorsal scan of a left wrist joint. This was the only joint region in which a colour fraction of $>$50% was measured. Table 4 compares the three grading systems.

The mean subjectively estimated colour fraction was higher than the mean measured colour fraction. The majority of joints were scored as grade 2 by the established classification system. Scores related to the colour fraction could provide a more even distribution. Table 5 describes ranges of subjectively estimated and measured colour

FIG. 1 Measurement of power Doppler fraction over 3 s




The highest peak represents the fraction described in this study. The corresponding US image is important for excluding artefacts.

fractions that would generate an even distribution of grades for the investigated RA patient population.

In six patients, 12 dorsal MCP regions and 3 dorsal wrist regions were evaluated by a second sonographer who was blinded to the results of the first sonographer. The analysis after the exercise found that both sonographers agreed in all joint regions in terms of the presence or absence of power Doppler signals in these joints. Sonographer 2 described a semi-quantitative grade 2 in two joint regions previously scored as grades 1 and 3 in five joint regions previously scored as grade 2. Sonographer 2 subjectively estimated and measured colour fractions of 45% (minimum–maximum 15–80%) and 32% (7–60%) in joints in which sonographer 1 had estimated and measured colour fractions of 26% (2–70%) and 16% (3–52%). Reliability was perfect for the presence or absence of colour signals. In terms of assignment to one of the three Doppler grades in joints exhibiting Doppler signals (>0/category 1–3 data), reliability was fair for the semi-quantitative grading system (weighted $\kappa$ 0.29), the estimated quantitative grading system (weighted $\kappa$ 0.37) and the quantitative computerized measurement grading system (weighted $\kappa$ 0.23). Reliability improved when comparing percentage measurements instead of grades. The weighted $\kappa$ was 0.57 for the estimated quantitative grading and 0.50 for the quantitative computerized measurement grading system.

It takes 5–10 min to bilaterally examine wrists and MCP joints with grey-scale and Doppler US, including documentation. Computerized Doppler quantification analysis takes an additional 3 min per joint with detectible synovial perfusion. On average, computerized analysis prolonged the examination time by 14 min, as each patient had an average of 4.7 joints with detectible synovial perfusion.

Table 2 Number of power Doppler–positive joint regions

| Joint region | Dorsal right | Dorsal left | Volar right | Volar left |
|---|---|---|---|---|
| Wrist | 5/23/2 | 4/23/2 | 4/1/0 | 1/3/0 |
| MCP 1 | 2/11/0 | 0/6/0 | 1/0/0 | 0/1/0 |
| MCP 2 | 5/17/2 | 6/8/1 | 1/2/0 | 0/0/0 |
| MCP 3 | 5/12/0 | 1/14/2 | 0/0/0 | 1/0/0 |
| MCP 4 | 0/5/1 | 2/5/0 | 0/0/0 | 0/0/0 |
| MCP 5 | 0/3/1 | 1/4/0 | 2/1/0 | 1/0/0 |
| Total | 17/71/6 | 14/60/5 | 8/4/0 | 3/4/0 |

First number indicates grade 1, second number grade 2 and third number grade 3 according to the Szkudlarek semi-quantitative grading system [12] modified by Naredo *et al*. [13].

## Discussion

The semi-quantitative grading system by Szkudlarek modified by Naredo for Doppler signals in synovitis is widely used. The present study shows that grades 1–3 are not evenly distributed in a population of consecutive patients with active RA. Only 5.7% of joint regions were scored as grade 3, subjectively estimating a colour fraction of >50%. In fact, measurement showed only one joint region with a colour fraction of >50% (52%). The colour fraction was overestimated by the sonographers. The measured colour fraction was lower than the estimated colour fraction. A quantitative grading system that relies only on estimated or measured colour fractions would allow more clearly defined cut-offs and a more even distribution of grades.

A general problem for all synovial Doppler US grading systems is the varying sensitivity of different US equipment and variable machine settings [39]. Strictly speaking, the suggested limits for the grades in this study account for only the equipment used with the mentioned machine settings and the current software. The colour fraction was lower than expected. The sensitivity for detecting Doppler signals may be higher with other US equipment or use of the colour Doppler mode with the equipment used in this study. Nevertheless, the present study was conducted with the newest equipment from the manufacturer. Thus the Doppler sensitivity in this study is probably higher than for most US examinations in clinical practice. As seen in previous studies, sensitivity for Doppler signals is higher in dorsal views than in volar scans [40].

Table 3 Mean estimated and measured colour fraction per joint region in power Doppler–positive joints

| Joint region | Dorsal right | | Dorsal left | | Volar right | | Volar left | |
|---|---|---|---|---|---|---|---|---|
| | Estimated, % | Measured, % | Estimated, % | Measured, % | Estimated, % | Measured, % | Estimated, % | Measured, % |
| Wrist | 18 | 11 | 20 | 10 | 12 | 6 | 13 | 7 |
| MCP 1 | 23 | 14 | 22 | 16 | 3 | 2 | 30 | 27 |
| MCP 2 | 21 | 14 | 18 | 10 | 13 | 5 | N/A | N/A |
| MCP 3 | 18 | 12 | 25 | 14 | N/A | N/A | 2 | 2 |
| MCP 4 | 27 | 17 | 20 | 13 | N/A | N/A | N/A | N/A |
| MCP 5 | 28 | 13 | 17 | 12 | 9 | 5 | 3 | 2 |
| Total | 21 | 13 | 21 | 12 | 10 | 5 | 12 | 8 |

N/A: not applicable, as in none of the patients were Doppler signals detected in these joint regions.

**Table 4** Estimated and measured colour fraction related to the conventional semi-quantitative grading system

| Semi-quantitative classification | Joint regions, *n* (%) | Colour fraction estimated, mean (s.d.), % | Colour fraction measured, mean (s.d.), % |
|---|---|---|---|
| Grade 1 | 42 (21.9) | 5.3 (3.6) | 3.9 (2.4) |
| Grade 2 | 139 (72.4) | 21.0 (10.5) | 12.6 (6.9) |
| Grade 3 | 11 (5.7) | 60.9 (6.6) | 29.7 (10.1) |
| Grades 1–3 | 192 | 19.7 (15.2) | 11.7 (8.6) |

**Table 5** Grading system with equal distribution of grades 1–3 for the investigated patient population

| | Estimated colour fraction, % | Joint regions | Measured colour fraction, % | Joint regions |
|---|---|---|---|---|
| Grade 1 | >0–10 | 68 | >0–6 | 57 |
| Grade 2 | >10–25 | 70 | >6–12 | 69 |
| Grade 3 | >25 | 54 | >12 | 66 |

Quantitative grading according to the estimated colour fraction could replace the previously used score in the future in order to get a better distribution of grades. As an alternative to a grading system of four categories (grades 0–3), the estimated or measured maximum colour fraction could be indicated. The sensitivity to change is likely to be superior to the old grading system because it allows a more even distribution of colour grades. This is particularly the case when the same equipment with the same software and settings is used for follow-up examinations. Further studies, particularly follow-up studies with treatment, will be needed to evaluate quantitative grading systems.

Computerized determination of the colour fraction appears to be more objective than estimation of the colour fraction. It has been shown that the reliability of computerized determination is higher for image interpretation [36]. Further studies are needed to determine if this also refers to image acquisition. Computerized determination of the colour fraction may gain a place particularly in trials and in selected institutions, but it is more time consuming than subjective estimation of colour signals. It is currently not yet widely available, but it will gain greater recognition in the future. Costs will arise mainly due to the increase in examination time rather than for the acquisition of technology. Quantitative scores might have a higher sensitivity to change [41]. Further follow-up studies are warranted to determine how much the sensitivity to change will be increased when applying the proposed quantitative scores.

The limitations of the study include the fact that the sonographers were not completely blinded to the clinical appearance of the patients. This is a general problem in US studies. The study population for the interrater analysis was small. Studies with a greater number of patients and more sonographers are needed to determine the use of new scoring systems for synovial vascularity.

In summary, computerized quantification of synovial perfusion in active RA results in lower colour fractions. Quantitative limit values for the respective grades may achieve a more even distribution than the established semi-quantitative score.

In conclusion, this study aims to open a discussion on the currently used semi-quantitative scoring system for synovial US Doppler signals. This score may need to be revised. Alternatively, new quantitative scores that rely on subjective estimation or electronic measurement of the maximum colour fraction in the synovium may replace semi-quantitative scores, with the aim of providing a more even distribution of grades and increasing the sensitivity to change, particularly for joints that have been classified as grade 2 by the old grading system. A new quantitative score would probably increase the sensitivity to change and thus improve the monitoring of treatment response in future trials.

## Acknowledgements

We thank Bernd Schicke, Tumorzentrum Berlin, Department of Statistics, Berlin, Germany, and Andrew Hutchings, London School of Hygiene & Tropical Medicine, London, UK, for statistical support.

*Funding*: The equipment for this study was provided by GE Healthcare Systems, Wauwatosa, WI, USA.

*Disclosure statement*: M.B. has received research grants from GE Healthcare and Esaote Germany. W.A.S. has received research grants from GE Healthcare, Esaote Germany and Siemens Medical Solutions. All other authors have declared no conflicts of interest.

## References

1 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.

2 Brown AK, Conaghan PG, Karim Z *et al*. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum 2008;58:2958–67.

3 Fukae J, Isobe M, Kitano A *et al*. Positive synovial vascularity in patients with low disease activity indicates smouldering inflammation leading to joint damage in rheumatoid arthritis: time-integrated joint inflammation estimated by synovial vascularity in each finger joint. Rheumatology 2013;52:523–8.

4 Schmidt WA, Backhaus M. What the practicing rheumatologist needs to know about the technical fundamentals of ultrasonography. Best Pract Res Clin Rheumatol 2008;22:981–99.

5 Newman JS, Laing TJ, McCarthy CJ *et al*. Power Doppler sonography of synovitis: assessment of therapeutic response: preliminary observations. Radiology 1996;198:582–4.

6 Schmidt WA, Völker L, Zacher J *et al*. Colour Doppler ultrasonography to detect pannus in knee-joint synovitis. Clin Exp Rheumatol 2000;18:439–44.

7 Walther M, Harms H, Krenn V *et al*. 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.

8 Weidekamm C, Köller M, Weber M, Kainberger F. Diagnostic value of high-resolution B-mode and Doppler sonography for imaging of hand and finger joints in rheumatoid arthritis. Arthritis Rheum 2003;48:325–33.

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

10 Carotti M, Salaffi F, Manganelli P *et al*. Power Doppler sonography in the assessment of synovial tissue of the knee joint in rheumatoid arthritis: a preliminary experience. Ann Rheum Dis 2002;61:877–82.

11 Hau M, Schultz H, Tony HP *et al*. Evaluation of pannus and vascularization of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis by high-resolution ultrasound (multidimensional linear array). Arthritis Rheum 1999;42:2303–8.

12 Szkudlarek M, Court-Payen M, Jacobsen S *et al*. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum 2003;48:955–62.

13 Naredo E, Rodríguez M, Campos C *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.

14 Naredo E, Bonilla G, Gamero F *et al*. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. Ann Rheum Dis 2005;64:375–81.

15 Koski JM, Saarakkala S, Helle M *et al*. Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. Ann Rheum Dis 2006;65:1590–5.

16 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.

17 Koski JM, Saarakkala S, Helle M *et al*. Assessing the intra- and inter-reader reliability of dynamic ultrasound images in power Doppler ultrasonography. Ann Rheum Dis 2006;65:1658–60.

18 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.

19 Bruyn GA, Naredo E, Möller I *et al*. Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis. Ann Rheum Dis 2009;68:357–61.

20 Cheung PP, Dougados M, Gossec L. Reliability of ultrasonography to detect synovitis in rheumatoid arthritis: a systematic literature review of 35 studies (1,415 patients). Arthritis Care Res 2010;62:323–34.

21 Bruyn GA, Pineda C, Hernandez-Diaz C *et al*. Validity of ultrasonography and measures of adult shoulder function and reliability of ultrasonography in detecting shoulder synovitis in patients with rheumatoid arthritis using magnetic resonance imaging as a gold standard. Arthritis Care Res 2010;62:1079–86.

22 Vlad V, Berghea F, Iagnocco A *et al*. Inter & intra-observer reliability of grading ultrasound videoclips with hand pathology in rheumatoid arthritis by using non-sophisticated Internet tools (LUMINA study). Med Ultrason 2014;16:32–6.

23 Backhaus TM, Ohrndorf S, Kellner H *et al*. The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. Ann Rheum Dis 2013;72:1163–9.

24 Naredo E, Möller 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.

25 Dougados M, Jousse-Joulin S, Mistretta F *et al*. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. Ann Rheum Dis 2010;69:828–33.

26 Hammer HB, Sveinsson M, Kongtorp AK, Kvien TK. A 78-joints ultrasonographic assessment is associated with clinical assessments and is highly responsive to improvement in a longitudinal study of patients with rheumatoid arthritis starting adalimumab treatment. Ann Rheum Dis 2010;69:1349–51.

27 Seymour MW, Kelly S, Beals CR *et al*. Ultrasound of metacarpophalangeal joints is a sensitive and reliable endpoint for drug therapies in rheumatoid arthritis: results of a randomized, two-center placebo-controlled study. Arthritis Res Ther 2012;14:R198.

28 Sakellariou G, Scirè CA, Verstappen SM, Montecucco C, Caporali R. In patients with early rheumatoid arthritis, the new ACR/EULAR definition of remission identifies patients

with persistent absence of functional disability and suppression of ultrasonographic synovitis. Ann Rheum Dis 2013;72:245–9.

29 Ohrndorf S, Backhaus M. Advances in sonographic scoring of rheumatoid arthritis. Ann Rheum Dis 2013; 72(Suppl 2):ii69–75.

30 Zufferey P, Brulhart L, Tamborrini G *et al*. Ultrasound evaluation of synovitis in RA: correlation with clinical disease activity and sensitivity to change in an observational cohort study. Joint Bone Spine 2014;81: 222–7.

31 Iagnocco A, Naredo E, Wakefield R *et al*. Responsiveness in rheumatoid arthritis. a report from the OMERACT 11 ultrasound workshop. J Rheumatol 2014;41:379–82.

32 Rakieh C, Nam JL, Hunt L *et al*. Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. Ann Rheum Dis 2014 Apr 12, doi: 10.1136/annrheumdis-2014-205227 [Epub ahead of print].

33 Reiche BE, Ohrndorf S, Feist E *et al*. Usefulness of power Doppler ultrasound for prediction of re-therapy with rituximab in rheumatoid arthritis: a prospective study of longstanding rheumatoid arthritis patients. Arthritis Care Res 2014;66:204–16.

34 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 2014;66:19–26.

35 Saleem B, Brown AK, Quinn M *et al*. Can flare be predicted in DMARD treated RA patients in remission, and is it important? A cohort study. Ann Rheum Dis 2012;71:1316–21.

36 Terslev L, Ellegaard K, Christensen R *et al*. Head-to-head comparison of quantitative and semi-quantitative ultrasound scoring systems for rheumatoid arthritis: reliability, agreement and construct validity. Rheumatology 2012;51:2034–8.

37 Ohrndorf S, Halbauer B, Martus P *et al*. Detailed joint region analysis of the 7-joint ultrasound score: evaluation of an arthritis patient cohort over one year. Int J Rheumatol 2013;2013:493848.

38 Aletaha D, Neogi T, Silman AJ *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

39 Ten Cate DF, Luime JJ, van der Ven M *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.

40 Witt MN, Mueller F, Weinert P *et al*. Ultrasound of synovitis in rheumatoid arthritis: advantages of the dorsal over the palmar approach to finger joints. J Rheumatol 2014;41:422–8.

41 Ellegaard K, Terslev L, Christensen R *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.