Agreement between semiquantitative and quantitative Doppler scoring systems for the assessment of synovial pathological vascularisation in rheumatoid arthritis

E. Naredo\textsuperscript{1,2}, I.M. Rutigliano\textsuperscript{3}, I. Janta\textsuperscript{1}, C. Scirocco\textsuperscript{3}, L. Valor\textsuperscript{4}, F. Figus\textsuperscript{4}, L. Carreño\textsuperscript{1}, J. Garrido\textsuperscript{5}, A. Iagnocco\textsuperscript{6}

\textsuperscript{1}Department of Rheumatology, Hospital General Universitario Gregorio Marañón, Madrid, Spain; \textsuperscript{2}Department of Rheumatology, Joint and Bone Research Unit, Hospital Universitario Fundación Jiménez Díaz, Autonoma University, Madrid, Spain; \textsuperscript{3}Dipartimento Medicina Interna e Specialità Mediche, Sapienza Università di Roma, Italy; \textsuperscript{4}Dipartimento Reumatologia, Università degli Studi di Cagliari, Italy; \textsuperscript{5}Department of Social Psychology and Methodology, Faculty of Psychology, Autonoma University of Madrid, Spain; \textsuperscript{6}Dipartimento Scienze Cliniche e Biologiche, Università degli Studi di Torino, Turin, Italy.

Abstract

Objective

To compare colour Doppler (CD) versus power Doppler (PD) semiquantitative and quantitative scoring of synovial vascularisation and to evaluate the relationship between semiquantitative and quantitative scores in patients with rheumatoid arthritis (RA).

Methods

One hundred RA patients underwent B-mode, PD, and CD assessments of 12 joints at two European centres. Each joint with synovial hypertrophy (SH) detected on B-mode was semiquantitatively scored (0-3) for PD (SPD score) and CD (SCD score) synovial signal. PD and CD synovial signal were also quantitatively scored (0-100%) (QPD and QCD scores, respectively) using a software integrated in the US equipment for counting the colour fraction.

Results

We found SH in 184 joints. SPD and SCD agreed in 92.3% (95%CI: 88.4; 96.2%) of paired scores, with Kendall rank correlation coefficient \(\tau=0.95\). QPD and QCD scores were highly correlated (Pearson’s coefficient=0.70) but Bland-Altman plot showed insufficient agreement, being the QCD scores systematically slightly higher than the QPD scores. The comparison of mean values of QPD and QCD between scores of SPD and SCD, respectively, showed significant differences between grade 0 and grade 1 (\(p<0.001\)), and grade 2 and grade 3 (\(p=0.042\) and \(p=0.007\), respectively) but not between grade 1 and 2 (\(p=0.154\) and \(p=0.150\), respectively).

Conclusion

The SPD and SCD scores were concordant and the QPD and QCD scores highly correlated but were not concordant. There was an overlap between SPD and SCD mild and moderate scores regarding QPD and QCD scores.

Key words

ultrasound, Doppler, rheumatoid arthritis, synovitis, semiquantitative and quantitative scoring systems
Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by intra-articular and periarticular synovial inflammation (i.e. synovial proliferation and angiogenesis) (1). Synovitis can damage the cartilage, bones, capsule, tendons, and ligaments with consequent joint deformities and severe joint function impairment.

Within the last decade, technological improvements in ultrasound (US) image resolution of musculoskeletal (MS) structures have led to an increasingly important role of this imaging tool in the evaluation and monitoring of patients with RA and other inflammatory arthritides based mainly on its higher capacity in detecting synovitis, when compared to clinical examination (2-4). MSUS is a routinely available, non-invasive, repeatable as many times as required at the time of consultation, and relatively inexpensive bedside technique with high patient acceptability. Colour Doppler (CD) and power Doppler (PD) techniques are able to detect synovial flow, which is an indirect sign of inflammatory activity (5-8). The presence of synovial Doppler signal has demonstrated strong predictive value in relation to structural damage progression and disease flare in both active and remission RA patients (9-13). The detection of synovial hypervascularisation is therefore of particular importance in the rheumatological evaluation and assessment of RA patients.

To the best of our knowledge only a few studies have compared different Doppler modalities (i.e. CD vs. PD) and different scoring systems (semiquantitative vs. quantitative) for synovial vascularity detected by Doppler US (14-16). The objectives of this cross-sectional observational study conducted in RA patients were the following: 1. to compare semiquantitative scoring of synovial vascularity by CD versus PD imaging; 2. to compare quantitative scoring of synovial vascularity by CD versus PD; 3. to evaluate the relationship between semiquantitative versus quantitative scoring of synovial vascularity by CD and PD.

Methods

Patients

One hundred patients [86 women, 14 men; mean (SD, range) age, 59.3 (13.4, 24–87) years] who fulfilled the American College of Rheumatology 1987 diagnostic criteria for RA (17) were consecutively included in a period of 3 years (17). They were recruited from the Rheumatology Departments of two European centres (Hospital Universitario Gregorio Marañón, Madrid, Spain, and Policlinico Umberto I - Sapienza Università di Roma, Rome, Italy).

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee of the Hospital General Universitario Gregorio Marañón, Madrid, Spain (the number of the study given by Ethics and Research Committee was 331/13) and the Policlinico Umberto I - Sapienza Università di Roma, Rome, Italy (the number of the study given by Ethics and Research Committee was 2965/13). Informed consent was obtained from all patients before study enrolment.

Clinical assessment

At each centre, all patients were evaluated by a local experienced rheumatologist, including the standard care clinical and laboratory assessments for RA patients (i.e. demographics, disease duration, Disease Activity Score for 28 joints (DAS28), presence of rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPA), and presence of radiographic erosions. The treatment taken by the patients was also registered.

Ultrasound assessment

The US assessments were performed by two rheumatologists, one at each centre (EN at Madrid and AI at Rome), who have who have respectively 20 and 32 years of experience in MSUS of experience in MSUS. They were blinded to clinical, laboratory and radiographic data. All patients underwent B-mode, PD, and CD assessments of 12 joints (18). These joints were as follows: bilateral elbow (i.e. anterior and posterior recesses), wrist (i.e. radiocarpal, midcarpal, and distal radioulnar
Intraarticular tissue that is nondisplaceable (NDACT) as the presence of abnormal hypoechoic material was assessed. Each joint was semiquantitatively scored for B-mode synovitis activity (20) according to the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) as the presence of abnormal hypoechoic material. This scoring system included a range of grades from 0 (absent) to 4 (marked), with 1, 2, and 3 representing moderate, high, and very high activity, respectively.

The US examinations were performed with a real-time scanner (LOGIQ E9, GE Medical Systems Ultrasound and Primary Care Diagnostics, LLC, Wauwatosa, WI, USA) equipped with a multifrequency linear matrix array transducer (ML6-15 MHz) used for elbow, wrist, knee, and ankle and a multifrequency linear hockey-stick transducer (L8–18 MHz) used for MCP joints. B-mode, PD, and CD settings were optimised for the different assessed joints before the study and standardised for the whole study (B-mode: frequency 13–18 MHz; gain 45–51. PD: frequency 6.3–12 MHz; gain 14–18; PRF 0.4–0.8. CD: frequency 6.3–11.9; gain 18–20; PRF 0.8).

Each joint was semiquantitatively scored (0–3) for SH, SPD, and SCD. SH was defined according to the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) as the presence of abnormal hypoechoic material. SH was scored for each joint as absolute frequencies and percentages. SH was found in 184 joints (15%) and synovial effusion in 133 joints (11%).

US findings
In total, we assessed 1,200 joints. We found SH in 184 joints (15%) and synovial effusion in 133 joints (11%). The distribution of SH, SPD, and SCD scores are shown in Table I. The mean (SD, range) MaxCF was 11.22% (14.14, 0.00–86.40) for PD and 13.81% (16.32, 0.00–95.60) for CD.

Agreement between power Doppler and colour Doppler
The agreement between PD and CD was studied at two levels: quantitative, comparing QPD and QCD findings, and semiquantitative, comparing SPD and SCD findings. At quantitative level, QPD and QCD showed a linear relationship with a Pearson coefficient of 0.842 (p<0.001), but the scatterplot showed a predominance of dots below the identity line (zero intercept and unit slope). The Bland-Altman showed an average of differences (QPD - QCD, against the average of these measurements (QPD + QCD)/2) of 0.15 with 95% confidence interval (Fig. 1). This bias was due to QCD values were slightly higher than their paired QPD values. The limits of agreement were -0.187 and 0.136 including between them 94.4% of differences. Differences inside limits of agreement represented up to 14.3% of mean QPD, which was above desirable.

At semiquantitative level, agreement between SPD and SCD was obtained in 92.3% (95% CI: 88.4; 96.2%) of paired scores, with discrepancies in 14 out of 184 pairs (7.7%). Kendall rank correlation coefficient tau-b was 0.75 (p<0.001). Additionally, the marginal homogeneity test did not find any significant differences between marginal distribution of SPD and SCD scores (p=0.565). The crosstab (Table II) shows that the 14 discrepancies between SPD and SCD scores were treated with synthetic disease-modifying anti-rheumatic drugs (DMARDs), 17% (17%) with biologic DMARDs, and 28% (28%) with synthetic and biologic DMARDs. Six (6%) patients were not receiving DMARDs. Forty-one (41%) patients were taking oral prednisone. Results

Demographics and RA characteristics
The mean (SD, range) disease duration was 10.3 (8.0, 0.5–40) years. The mean (SD, range) DAS28 was 3.94 (1.15, 1.9–7.19). RF was positive in 65 patients and ACPA in 57 patients. Sixty-nine patients showed radiographic erosions. Forty-nine (49%) patients were taking oral prednisone.
always between consecutive scores. These results suggested that agreement between SPD and SCD scores was adequate.

The progressivity of the semiquantitative variable was tested by one-way ANOVA with linear trend test. For both methods, one-way ANOVA demonstrated significant differences of means in QPD and QCD between SPD and SCD scores (p<0.001 in both). Moreover, a linear growing trend showing an increase of mean QPD and QCD as scores of SPD and SCD progressed was demonstrated (p<0.001 in both).

Table III displays the mean values of the QPD scores and QCD scores for each SPD score and SCD score, respectively. For PD scores, Games-Howell comparisons between the different grades were as follows; SPD=0 vs. SPD=1, p<0.001; SPD=0 vs. SPD=2, p=0.001; SPD=0 vs. SPD=3, p=0.001; SPD=1 vs. SPD=2, p=0.154; SPD=1 vs. SPD=3, p=0.012; and SPD=2 vs. SPD=3, p=0.042. There was an overlapping of QPD scores between SPD scores 1 and 2. For CD scores, Games-Howell comparisons between the different grades were the following: SCD=0 vs. SCD=1, p=0.001; SCD=0 vs. SCD=2, p<0.001; SCD=0 vs. SCD=3, p=0.001; SCD=1 vs. SCD=2, p=0.150; SCD=1 vs. SCD=3, p=0.003; and SCD=2 vs. SCD=3, p=0.007. Again, there was an overlapping of QCD scores between SCD scores 1 and 2.

Comparison of mean values of QPD and QCD between scores of SPD and SCD, respectively is shown in Figures 2 and 3. We used parametric test in order to check linear trend in means. Nevertheless, non-parametric test of Kruskal-Wallis offered similar results for QPD and QCD. Non-parametric tests confirm the findings with parametric tests. (Table IV).

Discussion
In recent years, Doppler modalities have been increasingly used to assess joint synovitis in arthritides and different scoring systems have been proposed to grade synovial perfusion in RA (21). The present study has demonstrated that semiquantitative scores assessed by CD and PD are concordant with no significant differences between their distribution. Thus, the semiquantitative scores appear to be independent by the Doppler modality that is used, being both similarly able to detect pathologic flow within the synovial tissue of RA patients. This aspect is of particular value for ultrasonographers in their choice of the semiquantitative modality to use. The evaluation of the distribution of quantitative scores showed that they were highly correlated but not concordant. Indeed, the QCD scores were systematically slightly higher than the QPD scores, thus showing that CD detects more signal than PD when quan-
In addition, our study demonstrated consistency between SPD and SCD moderate and severe scores and QPD and QCD scores and an overlap between SPD and SCD mild and moderate scores regarding QPD and QCD scores. Both assessment systems were consistent for absence of synovial Doppler signal. Quantitative and semiquantitative assessments are both valuable in detecting pathological flow, however, the intrinsic characteristics of them differ in terms of the modality used for signal scoring. In both CD and PD modes, semiquantitative and quantitative evaluations analyse and score the area of signal in grades 2 and 3. Conversely, grade 1 is scored on the basis of the number of signals by the semiquantitative scoring system and, instead, the area of signal is taken into account by QPD evaluation. This discrepancy in the assessment of flow may play a role in the results of the grading, making the semiquantitative evaluation different than the quantitative methods which is always based on the area of signal in all scores. However, because quantitative assessment is more time consuming than semiquantitative evaluation, its widespread use in the clinical practice is limited and it may gain a place particularly in research. Nevertheless,
the mentioned overlap between both scoring system for mild and moderate scores need further attention regarding a potential synovial Doppler cut-off point with either diagnostic or prognostic purpose.

CD and PD are the two most commonly used Doppler modes to assess synovial hypervascularisation in rheumatic diseases (22). They are the most studied and applied because they allow the simultaneous visualisation of grey-scale and Doppler findings, providing information on the exact anatomic distribution and on the entity of the blood flow (22, 23). In the most commonly used high-end US equipment, CD and PD are considered to have equal sensitivity in detecting pathological vascularisation in arthritides (16, 23, 24). This is mainly due to the advancing US technology and the correlated increased sensitivity of recent machines in detecting Doppler signal (23). By using both modes, either semiquantitative or quantitative scoring systems can be applied to grade the severity of synovial perfusion, however, there is not yet consensus on which of them might better reflect the pathological vascularisation in RA, although semiquantitative assessment seems to be more feasible clinical practice (23, 24).

Our study allowed us to assess a high number of joints, with a combination of small and large joints. However, a limitation of our research need to be addressed as we tested Doppler modes only in machines of a single manufacturer, while it would have been interesting to analyse the discrepancies also in other equipment. Nonetheless, this study was conducted with the highest end and newest equipment which aspect should minimise this bias.

In summary, this study, aimed to compare CD versus PD semiquantitative and quantitative scoring of synovial vascularisation in RA and to evaluate the relationship between scoring systems, demonstrated that semiquantitative PD and CD scores were concordant. Quantitative PD and CD scores were consistent with the respective semiquantitative scores for scores moderate and severe. However, PD and CD mild and moderate scores showed an overlapping of quantitative PD and CD values, respectively. Further longitudinal studies comparing the sensitivity to change of semiquantitative versus quantitative PD and CD scores are warranted.

References