

Doubtful swelling on clinical examination reflects synovitis in rheumatoid arthritis

Peter Mandl¹, Paul Studenic, Gabriela Supp, Martina Durechova, Stefanie Haider, Michaela Lehner, Tanja Stamm, Josef S. Smolen and Daniel Aletaha

Ther Adv Musculoskel Dis

2020, Vol. 12: 1–9

DOI: 10.1177/
1759720X20933489

© The Author(s), 2020.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract

Aims: The aim of this study was to evaluate the implication of doubtful joint swelling on clinical examination with respect to objective markers of synovitis by ultrasound (US) in patients with rheumatoid arthritis (RA).

Methods: Two independent observers performed a modified 28 swollen joint assessment (28SJC), in which joints could be graded as either definitely swollen, non-swollen, or doubtfully swollen. Two examiners blinded to clinical information performed US assessment of the hands. We performed descriptive statistics and models to analyse the links between clinical assessment and objective markers of inflammation.

Results: A total of 1204 joints were evaluated in 43 RA patients; 93% (40/43) of patients had ≥ 1 joint with doubtful swelling (range: 0–4/patient). Inter-reader reliability for the modified 28SJC was good (0.74). Generally, both grey scale (GS) and power Doppler (PD) discriminated across not swollen, doubtful, and swollen joints. GS signals discriminated better than PD between doubtful swelling and no swelling [odds ratio (OR) for GS: 5.2; 95% confidence interval (CI) 1.2–23.3 versus OR for PD 1.7; 95% CI 0.2–13.0], whereas PD discriminated better than GS between swelling and doubtful swelling [OR for PD: 28.7; 95% CI 3.6–228.2 versus GS: 1.7; 95% CI 0.3–8.4]. Joint osteophytes did not increase the degree of doubtfulness.

Conclusion: Clinical doubt in the assessment of joint swelling constitutes an intermediate state between unequivocal swelling and the lack thereof also regarding the objectively quantified level of inflammation. In order to increase sensitivity for joint inflammation, the historical clinical approach of considering doubtful swelling the absence of swelling should be revisited to interpret clinical doubtfulness as an indication of swelling.

Keywords: clinical joint examination, doubtfulness, rheumatoid arthritis, ultrasound

Received: 21 January 2020; revised manuscript accepted: 19 May 2020.

Introduction

Clinical evaluation of joint involvement, specifically the assessment of tender and swollen joints, remains the hallmark of both the diagnosis as well as the monitoring of arthritis. Joint damage and impairment of physical function are the most important adverse outcomes of rheumatoid arthritis (RA), a consequence of clinical disease activity over time. Joint damage is particularly associated with swollen joint counts (SJC) and acute phase reactant levels as well as composite measures of disease activity that comprise these variables.^{1–3} Grading the extent of clinical joint swelling or tenderness does not

improve the performance of a score,⁴ and, therefore, joint swelling is recommended to be performed in an ungraded fashion.⁵ To facilitate assessment in clinical practice, it has been recommended that a joint should be classified as clinically swollen only if the swelling is beyond doubt.⁶ However, it may be difficult to evaluate and confirm joint swelling by clinical assessment in certain situations. In patients with pudgy or oedematous fingers, swelling due to synovitis may be difficult to distinguish from mere extra-articular soft-tissue changes on clinical examination. Indeed, a recent study found, that clinical joint swelling is less likely

Correspondence to:

Peter Mandl
Division of Rheumatology,
Department of
Rheumatology, Medical
University of Vienna,
18-20, Währinger Gürtel,
Vienna, 1180, Austria
[peter.mandl@
meduniwien.ac.at](mailto:peter.mandl@meduniwien.ac.at)

Paul Studenic
Gabriela Supp
Martina Durechova
Michaela Lehner
Daniel Aletaha
Department of
Rheumatology, Medical
University of Vienna,
Vienna, Austria

Stefanie Haider
Orthopedic Hospital
Speising, Institute of
Physical Medicine and
Orthopaedic Rehabilitation,
Vienna, Austria

Tanja Stamm
Center for Medical
Statistics, Informatics,
and Intelligent Systems,
Medical University of
Vienna, Vienna, Austria

Josef S. Smolen
Medical University of
Vienna, Vienna, Austria

to represent synovitis [using power Doppler (PD) ultrasound (US) as reference] in obese RA patients as compared with those with a normal body mass index (BMI).⁷ Further, certain joints may be more difficult to assess than others, for example, the shoulder or foot joints.⁸ These reasons may lead the examiner to 'doubt' clinical swelling of a joint and thus potentially underestimate the level of disease activity, in particular because the current standards adhere to the view of 'when in doubt indicate as non-swollen'.⁶ In the present study we granted the joint examiner the option to label a joint as doubtfully swollen and investigated whether such joints would show increased US signs of subclinical activity as compared with non-swollen joints, and less than joints that are clinically swollen beyond doubt.

Patients and methods

Patients

Inclusion criteria comprised fulfilment of the 2010 American College of Rheumatology/European League against Rheumatism (ACR/EULAR) classification criteria for RA,⁹ the presence of at least one clinically unequivocally swollen joint, and age of >18 years. Enrollment of patients was consecutive. We aimed to recruit a sample size of 43 patients, corresponding to 1204 joints. This provided a sufficient sample size to use grey-scale (GS) or PD signs of synovitis as outcome variables for assessment of the potential ambiguity of clinical joint swelling in a logistic regression analysis. The sample size calculation was based on an assumed sampling ratio of 0.1 between doubtfully swollen and non-swollen joints,¹⁰ and an estimated odds of 0.3 and 0.1, respectively, for the event of detecting PD.¹¹ The sample size calculation was based on the joint as the unit of analysis, and not the patient. In a sensitivity analysis, we addressed the impact of non-independence of the data units (i.e. the fact that multiple joints stem from the same patient) on the results, by randomly selecting one doubtful (by any examiner), one non-swollen and one swollen joint per individual patient. The ethics committee of the Medical University of Vienna approved the study, which was conducted according to the guidelines of the Declaration of Helsinki. All patients provided written consent upon inclusion into the study.

Assessments

Clinical assessment. Swollen joint counts were performed on the 28-joint count scale (28SJC)

(bilateral wrist, metacarpophalangeal joints 1–5, interphalangeal joint of the thumb, proximal interphalangeal joints 2–5, elbow, shoulder, and knee joints).^{7,12} In 43 patients with RA two independent biometricians, health professionals with more than 5 years of experience performing daily joint counts in patients with arthritides, who were blind with regard to the examination of their co-evaluator as well as with regard to sonographic data, performed a modified 28SJC by assigning the following labels to individual joints with regard to swelling: (a) definitely swollen; (b) doubtfully swollen and (c) definitely not swollen. Doubtful swelling was defined as a state in which the examiner was unable to clinically rule out or confirm synovial swelling. For the purpose of the primary analysis, a joint was labelled as doubtfully swollen when swelling was doubtful to both examiners.

US and radiographic assessment. Following clinical evaluation by the biometricians, a standardised US evaluation of the 43 patients was performed by a sonographer blinded to the clinical data (G.S.). A systematic, multiplanar US examination was carried out using both GS and PD with a real-time scanner (General Electric Logiq E9) with a multifrequency linear transducer (10–15 MHz). Both GS and PD examinations were recorded for each of the 28 joints assessed. The sonographic evaluation of each joint was carried out according to the standardised scanning technique described in the EULAR guidelines and included both dorsal and volar scans.¹³ Presence and absence of GS signs of synovitis and intra-articular PD signal as well as presence/absence of bone erosions and of osteophytes as defined by the OMERACT Working Group were recorded for each joint on a pre-designed form.¹⁴ Conventional anterior-posterior radiographs of the hands collected annually as part of the routine follow up of RA patients at our department were evaluated for the presence/absence of osteophytes using the Interphalangeal Osteoarthritis Radiographic Simplified (iOARS) score.¹⁵

Statistical analysis

Agreement between examiners on the modified 28 swollen joint assessment was assessed on joint level by weighted Fleiss kappa (agreement beyond chance quantified between 0 and 1). Intrareader reliability for the sonographer performing the US assessments for assessing synovitis using the same US machine and the same settings was found to

be good to excellent (0.665 and 0.972 for PD and GS US, respectively) in a previous study.¹⁶ To evaluate the association between clinical assessment of joint swelling and the presence of sonographic signs of synovitis, we applied logistic regression analyses: we utilized positivity for GS and PD as dependent variables, and doubtfully swollen *versus* non-swollen or swollen as explanatory variables, followed by chi-square test. To avoid retrieving mean results, we performed the same logistic regression analyses for each examiner individually.

The influence of demographic variables such as age, gender, disease duration and BMI as well as descriptive, clinical and laboratory characteristics: rheumatoid factor (RF), anti-citrullinated antibodies (ACPA), C-reactive protein (CRP), evaluators global assessment (EGA), erythrocyte sedimentation rate (ESR), fatigue (visual analogue scale 0–10), health assessment questionnaire (HAQ), pain (visual analogue scale 0–10), stiffness (visual analogue scale 0–10), patient global assessment (PGA), disease activity score with CRP or ESR (DAS28-CRP and DAS28-ESR), swollen joint count (SJC), tender joint count (TJC), simplified disease activity index (SDAI) and clinical disease activity index (CDAI) were then evaluated using linear regression. In order to see whether patients with high counts of doubtfully swollen joints differ from those with low counts, we divided them into tertiles based on the number of doubtfully swollen joints, and calculated trends over these tertiles of differences in descriptive, clinical and laboratory characteristics using the Jonckheere-Terpstra test.

We then examined the influence of potential osteoarthritis on ambiguity of joint assessment, by evaluating the number of osteophytes in each of the different swelling categories. Additionally, we reran the above-mentioned model omitting joints that showed signs of osteophytes by US.

In a sensitivity analysis, we calculated a mean numerical swelling status for each joint taking the evaluation of both examiners into account by calculating a score ranging from 0 to 2 for each joint (0 = non-swollen, 1 = doubtfully swollen, and 2 = definitely swollen) divide by 2 (for the two examiners). These scores ranged from 0 to 2 in steps of 0.5 and were used it in the logistic regression model. The dependent variables in this logistic regression analysis were GS signs of synovitis, PD signal and presence of osteophytes on US.

Data analyses were carried out using SPSS®, Version 25 (SPSS, Chicago, IL, USA) and STATA (StataCorp. 2017. Stata: Release 15. Statistical Software. StataCorp LLC, College Station, TX, USA). Significance level was 0.05.

Results

Frequency of doubtful joint swelling and respective agreement

A total of 1204 joints were evaluated in 43 RA patients (Table 1); 93% (40/43) of patients had ≥ 1 DSJ, with a maximum number of 4 DSJ/patient. Doubtfully swollen joints by one examiner were classified so by the other examiner in only 17%; in the majority of cases they were classified differently, and to a comparable proportion as non-swollen (45% of cases) and swollen joints (38%).

The joints that were most frequently found to be doubtfully swollen by at least one examiner were the wrist, proximal interphalangeal (PIP) 3, PIP 5 and metacarpophalangeal (MCP) 2 joints, with the elbow and the shoulder joints least commonly doubtful (Figure 1), but also least commonly affected by swelling. Inter-examiner reliability for the modified (allowing for doubtful swelling) 28 swollen joint assessment was good (0.74; 95% CI: 0.70–0.79).

Sonographic characterization of doubtfully swollen joints

The prevalence of sonographic signs of synovitis in joints according to swelling status is shown in Table 2. Based on logistic regression modelling, the comparative risk for sonographic signs of inflammation for the different states by clinical assessment was calculated.

Results are cross-tabulated and summarized for GS and PD findings in Figure 2. Generally, for both GS and PD, ORs to the left and below the diagonal are < 1 , whereas ORs to the right and above the diagonal are > 1 , indicating also an inflammatory continuity from no swelling to doubtful swelling to definite swelling. GS signs of synovitis discriminated better between no swelling and various grades of doubtful swelling as assessed by two observers: DSJ/NS or DSJ/DSJ *versus* NS/NS: OR 2.4 (95% CI: 1.2–4.9) and 5.2 (95% CI: 1.2–23.2) $p < 0.05$ for both, respectively (Figure 2A); than PD signal: OR 2.1 (95% CI: 1.0–4.5) $p = 0.056$;

Table 1. Descriptive characteristics of study patients ($n=43$).

Age, years	57.1 (13.2)
Female	90.5%
Disease duration, years	12.3 (10.5)
Rheumatoid factor, IU/ml	105.2 (141.5)
Rheumatoid factor positive in %	65%
Anti citrullinated antibodies, IU/ml	128.4 (128.8)
Anti citrullinated antibodies in %	65%
C-reactive protein, mg/dl	0.6 (0.8)
Erythrocyte sedimentation rate	30.3 (16.9)
HAQ	1 (2.1)
Patient Global Assessment, in mm	27.6 (26.4)
Evaluator Global Assessment, in mm	20.5 (14.6)
Pain, in mm	25.7 (23.9)
28 swollen joint count	3.6 (3.1)
28 tender joint count	2.1 (4.1)
Fatigue, in mm	22.8 (21.5)
DAS28	4.9 (6.6)
SDAI	11 (9.2)
CDAI	10.8 (8.7)
Joint space narrowing (Sharp van der Heijde score)	28 (25.9)
Erosions (Sharp van der Heijde score)	16.7 (20.7)
Total Sharp van der Heijde score	42.6 (44.4)

Values are displayed as means (SD), unless indicated otherwise.
CDAI, clinical disease activity index; DAS28, disease activity score 28; HAQ, health assessment questionnaire; SD, standard deviation; SDAI, simplified disease activity index; IU, international unit.

and 1.7 (95% CI: 0.2–13.0) $p=0.618$, respectively (Figure 2B). In contrast, in more active joints, PD signal discriminated better than GS, as for the comparison of SJ/SJ *versus* DSJ/SJ or DSJ/DSJ: OR for PD signal were as follows: 3.1 (95% CI: 1.4–6.7) and 28.7 (3.6–228.2) $p<0.01$ for both, respectively (Figure 2B); while the OR for GS signs of synovitis were: 0.9 (95% CI: 0.2–3.5) $p=0.878$ and 1.6 (95% CI: 0.3–8.4), $p=0.548$, respectively (Figure 2A). Evaluation of the association between doubtfulness and US findings on patient level confirmed the results of the analysis on joint level.

Sensitivity analyses

We performed two sensitivity analyses. First, we performed the above analysis separately for each examiner, which were confirmatory of the overall findings: when compared with non-swollen joints, doubtfully swollen joints were more likely to exhibit both GS signs of synovitis: OR examiner 1 was 5.4 (95% CI: 2.3–13.0) $p<0.001$; and OR examiner 2 was 3.4 (95% CI: 1.8–6.4) $p<0.001$; and PD signal: OR examiner 1 was 3.1 (95% CI: 1.6–6.1) $p<0.001$; and OR examiner 2 was 1.9 (95% CI: 1.0–3.7) $p<0.05$. Similarly, when compared with joints with doubtful swelling, joints with definite swelling were even more likely to exhibit PD signal: OR examiner 1 was 3.6 (95% CI: 1.8–7.4) $p<0.001$; and OR examiner 2 was 5.6 (95% CI: 2.8–11.5) $p<0.001$ (Table 3). GS signs of synovitis alone were found to be less discriminative between doubtful swelling and definite swelling: OR examiner 1 of 1.3 (95% CI: 0.5–3.6) $p=0.590$ and OR examiner 2 of 2.2 (95% CI: 0.9–5.0) $p=0.074$.

In another sensitivity analysis, we used a modified swelling metric (ranging from 0 to 2 in steps of 0.5), which integrated the rating of both examiners, and thus reduced noise by inter-rater variability. Joints with higher mean measurements were more likely to exhibit GS signs of synovitis: OR 3.3 (95% CI: 2.5–4.5) $p<0.001$; and PD signal: OR 3.8 (95% CI: 3.1–4.7) $p<0.001$ (Figure 3), indicating again that doubtful swelling is characterized by higher GS signs of synovitis and PD signal as compared with the absence of swelling and lower signal as compared with definitely swollen joints. Finally, we could not demonstrate any association between radiographic or sonographic osteophytes and doubtful swelling joints (labelled doubtful by one or both examiners) (Figure 3C; other data not shown).

Patient factors associated with doubtfulness

We found no correlation between age, gender or BMI and the presence of doubtful swelling. After dividing the patients into tertiles based on the number of doubtfully swollen joints, patients with doubtfully swollen joints tended to have a shorter disease duration ($p=0.045$) and symptom duration ($p=0.029$). Patients exhibiting a higher number of doubtfully swollen joints had significantly higher disease activity as indicated by the traditional SDAI and CDAI ($p=0.006$ and 0.008 respectively), but not by the DAS28-CRP or DAS28-ESR, driven mainly by a higher overall SJC28 ($p<0.001$) and

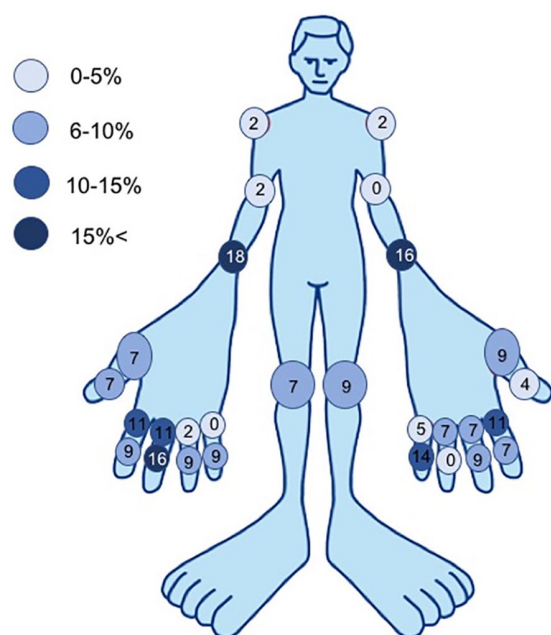


Figure 1. Frequency of doubtful joints. Numbers represent the percentage of patients ($n=44$), in which the respective joint was classified as doubtful by at least one observer; joints included in the 28JCS are shown.

28JCS, 28-joint count scale.

EGA ($p < 0.001$). We found no significant association between the number of doubtfully swollen joints and presence or levels of RF, ACPA, CRP, ESR, fatigue, HAQ, pain, PGA, stiffness or tender joint count (Supplemental Table S1).

Discussion

GS signs of synovitis and PD signal on US were more common in doubtfully swollen joints than non-swollen joints, and less common than in definitely swollen joints, suggesting that ‘doubtful’ swelling may indeed represent not only an intermediate state of clinical ambiguity, but also underlying inflammatory joint activity.

At the same time, the clinical state of ‘doubtfulness’ as a grey zone between clear presence and absence of joint swelling is very subjective. This is also reflected by the low agreement between examiners for doubtful swelling. They agreed on ‘doubtfulness’ in only 17% of the joints that were labelled as doubtfully swollen by any one of them, and the remaining 83% of these joints were split by the respective investigator about 50–50% into non-swollen and swollen joints,

indicating that these non-concordantly judged joints were at least highly controversial with respect to their swelling status. This is in line with previous observations,^{17–19} and in fact supports former conclusions on the appropriateness of the ungraded evaluation of joint swelling.⁵ However, when the individual assessments of the two examiners were compared separately to the US findings, results for both GS signs of synovitis and PD signal were confirmatory of the main analysis (which classified joints based on the concordant adjudication).

Of interest, our results using the modified swelling metric, suggest that synovitis as assessed by GS may be more sensitive in discriminating differences in inflammation on the lower end of the scale, whereas with higher levels of inflammation GS signals may not be able to increase further; for the increased vascularity as assessed by PD, it may be the opposite, with less discriminatory capacity in case of low-level inflammation, and better discrimination in more highly inflamed joints, in line with recent studies that imply that PD signal may be a better tool to detect active joint inflammation.^{20–22} This different range of discriminatory sensitivity is supported by comparison of panels A and B of Figure 2, and even more so by comparison of the slopes of the regression curves in Figure 3A and B.

The fact that patients with shorter disease duration and higher disease activity more often exhibited doubtful swelling suggests an association between ambiguity of swelling and more established disease. The number of doubtful joints was also higher in patients with a higher number of swollen joints,

Table 2. Prevalence of GS signs of synovitis, PD signal and osteophytes according to swelling status.

Observer ratings	GS synovitis % (n)	PD % (n)	Osteophytes % (n)
NSJ/NSJ	53.6 (501/934)	11.5 (107/934)	11.1 (104/934)
NSJ/DSJ	73.8 (31/42)	21.4 (9/42)	26.2 (11/42)
DSJ/DSJ	85.7 (12/14)	7.1 (1/14)	14.3 (2/14)
DSJ/SJ	91.7 (33/36)	41.7 (15/36)	16.7 (6/36)
SJ/SJ	90.8 (99/109)	68.8 (75/109)	9.2 (10/109)

DSJ/DSJ, joints rated as doubtful by both observers; DSJ/SJ, joints rated as doubtful by one observer and swollen by the other observer; GS, grey scale; NSJ/DSJ, joints rated as doubtful by one observer and non-swollen by the other observer; NSJ/NSJ, joints rated non-swollen by both observers; PD, power Doppler; SJ/SJ, joints rated swollen by both observers.

(A)	Comparator status: OR for GS (and 95% confidence interval) compared to reference status				
Reference status	NS/NS	DSJ/NS	DSJ/DSJ	DSJ/SJ	SJ/SJ
NS/NS	1	2.4 (1.2-4.9)*	5.2 (1.2-23.3)*	9.5 (2.9-31.2)***	8.6 (4.4-16.6)***
DSJ/NS	0.4 (0.2-0.8)*	1	2.1 (0.4-11.1)	3.9 (1.0-15.3)	3.5 (1.4-9.1)**
DSJ/DSJ	0.2 (0-0.9)*	0.5 (0.1-2.4)	1	1.8 (0.3-12.3)	1.7 (0.3-8.4)
DSJ/SJ	0.1 (0-0.3)***	0.3 (0.1-1.0)	0.5 (0.1-3.7)	1	0.9 (0.2-3.5)
SJ/SJ	0.1 (0-0.2)***	0.3 (0.1-0.7)**	0.6 (0.1-3.1)	1.1 (0.3-4.3)	1

(B)	Comparator status: OR for PD (and 95% confidence interval) compared to reference status				
Reference status	NS/NS	DSJ/NS	DSJ/DSJ	DSJ/SJ	SJ/SJ
NS/NS	1	2.1 (1.0-4.6)	1.7 (0.2-13.0)	5.5 (2.8-11.0)***	17.1 (10.8-26.8)***
DSJ/NS	0.5 (0.2-1.0)	1	0.3 (0-2.5)	2.6 (1.0-7.0)	8.1 (3.5-18.8)***
DSJ/DSJ	0.6 (0.1-4.6)	3.5 (0.4-30.8)	1	9.3 (1.1-78.9)*	28.7 (3.6-228.2)**
DSJ/SJ	0.2 (0.1-0.4)***	0.4 (0.1-1.0)	0.1 (0-0.9)*	1	3.1 (1.4-6.7)**
SJ/SJ	0.1 (0.1-0.1)***	0.1 (0.1-0.3)***	0 (0-0.2)**	0.3 (0.2-0.7)**	1

Figure 2. OR and 95% CI (in brackets) of joints calculated by logistic regression, for the detection of (A) GS signs of synovitis and (B) PD signal as compared with the reference swelling status. CI, confidence interval; DSJ/DSJ, joints rated as doubtful by both observers; DSJ/SJ, joints rated as doubtful by one observer and swollen by the other observer; GS, grey-scale; NSJ/DSJ, joints rated as doubtful by one observer and non-swollen by the other observer; NSJ/NSJ, joints rated non-swollen by both observers; OR, odds ratio; PD, power Doppler; SJ/SJ, joints rated swollen by both observers. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Table 3. OR (95% CI) for the detection of GS signs of synovitis and PD signal in relation to swelling status for each examiner.

		Examiner 1	Examiner 2
Grey-scale	DSJ versus NSJ	5.445 [2.284–12.981]***	3.371 [1.771–6.417]***
	SJ versus DSJ	1.317 [0.483–3.593]	2.152 [0.930–4.980]
Power Doppler	DSJ versus NSJ	3.127 [1.594–6.135]***	1.925 [1.013–3.658]*
	SJ versus DSJ	3.613 [1.757–7.429]***	5.675 [2.812–11.453]***

CI, confidence interval; DSJ, joint rated as doubtful; GS, grey scale; NSJ, joint rated non-swollen; OR, odds ratio; PD, power Doppler; SJ, joints rated as swollen.
* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

which is not surprising if one considers doubtful-ness to be an intermediary state of inflammatory activity. The fact that we demonstrated correlation between the number of doubtful joints and disease activity as assessed by the SDAI and CDAI, but not by DAS28 might again be explained by the lower weight of the SJC in the latter index. The lack of

association between the number of doubtful joints and many other patient and disease activity factors is likely related to the weaker association of these factors with inflammatory joint activity as such.

It is noteworthy to mention that the phenomenon of doubtful swelling may also occur during the

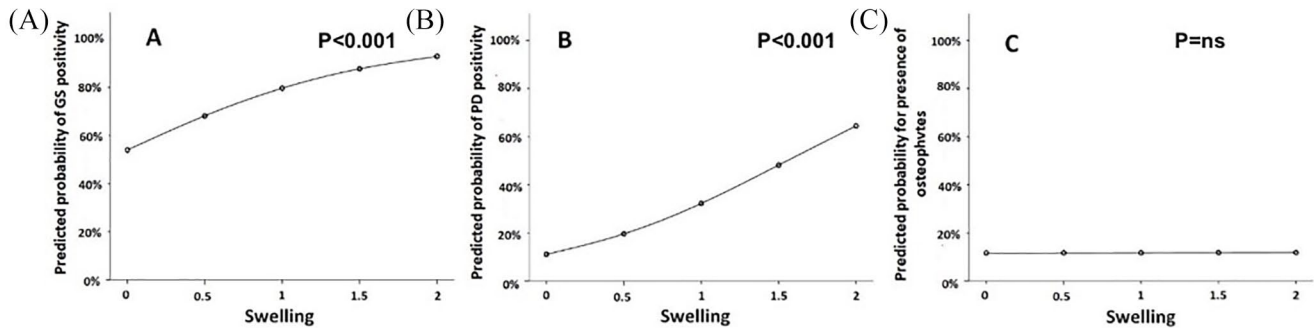


Figure 3. Predicted probability of GS signs of synovitis (A), PD signal (B), presence of osteophytes (C) on US in relation to mean swelling status (ranging from 0 to 2, with 0 non-swollen, 1 doubtfully swollen and 2 definitely swollen, taking both examiners into account as mean). GS, grey scale; ns, nonsignificant; PD, power Doppler; US, ultrasound.

clinical assessment of other rheumatic and musculoskeletal diseases such as psoriatic arthritis and should thus be further investigated in other conditions and cohorts. Performing a semiquantitative assessment of sonographic findings as well as other forms of imaging such as magnetic resonance imaging in future studies may allow us to further substantiate the state of doubtful joint swelling.

Although more than 1000 joints were assessed, enabling us to arrive at these conclusions, our study may be regarded as limited due to a relatively small sample size. In addition, the frequency of unequivocal DSJ sites was low; however, a number of sensitivity analyses and separate assessment of the two examiners revealed fully confirmatory results.

When designing the sonographic protocol for our study, our goal was to simulate as far as possible the clinical examination, which is why we choose binary grading performed by the sonographer during the examination rather than *post hoc* scoring of static images. While binary grading is less suitable as compared with semiquantitative scoring for monitoring treatment in longitudinal studies, our study was not designed to monitor therapeutic response and, in addition, previous studies have shown that binary grading is more reliable than semiquantitative scoring.^{23,24} However, the use of binary, rather than semiquantitative grading, and the difficulties in defining a threshold of GS and PD indicative of active synovitis, recently highlighted by a number of recent studies may be considered as a limitation of our study.^{25,26} In this study, we used US, a technique that has been shown to be more sensitive than clinical examination, not as a gold standard but as a surrogate for inflammation. The goal of the study

was to evaluate the ambiguity of clinical joint swelling, and not to draw conclusions on the use of US for monitoring disease activity. Furthermore, we focused exclusively on articular rather than extraarticular structures. Evaluation of extraarticular structures including tendons might have provided valuable information on clinical joint ambiguity.

Our data support the notion that doubtful joints on clinical examination represent an intermediate state between swelling and absence of swelling, when compared with more sensitive imaging techniques, such as sonography. In principal, this challenges the current logic of ‘when in doubt indicate as non-swollen’ with the potential conversion to ‘when in doubt indicate as swollen’.⁶ Changing this paradigm in clinical practice as stated above may lead to higher swollen joint counts and consequent categorization in higher disease activity states, as was the case in one of our earlier studies investigating multimodal disease activity indices utilising both sonographic and clinical data.²⁷ Such practice might have therapeutic consequences, and may run the risk of overtreatment. At the same time, the opposite (and current practice) runs the risk of undertreatment.

In summary, our data might indicate that the sensitivity of clinical joint assessment for inflammation can be increased when doubtful joints are considered as swollen and included into joint counts and composite scores.

Conflict of interest statement

P.M. reports grants and personal fees from AbbVie, BMS, Chugai, MSD, Janssen, Lilly, Novartis, Pfizer, Roche, outside the submitted

work; T.S. reports personal fees from AbbVie, Janssen, MSD, Novartis and Roche, outside the submitted work; J.S. received grants to his institution from Abbvie, AstraZeneca, Janssen, Lilly, Merck Sharpe & Dohme, Pfizer, and Roche and provided expert advice for, or had symposia speaking engagements with, AbbVie, Amgen, AstraZeneca, Astro, Bristol-Myers Squibb, Celgene, Celltrion, Chugai, Gilead, Glaxo, ILTOO Pharma, Janssen, Lilly, Merck Sharp & Dohme, Novartis- Sandoz, Pfizer, Roche, Samsung, Sanofi, and UCB; P.S., G.S., M.D., S.H., M.L. and D.A. have nothing to disclose.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Mandl Peter  <https://orcid.org/0000-0003-1526-4052>

Supplemental material

Supplemental material for this article is available online.

References

- Aletaha D and Smolen JS. Joint damage in rheumatoid arthritis progresses in remission according to the disease activity score in 28 joints and is driven by residual swollen joints. *Arthritis Rheum* 2011; 63: 3702–3711.
- Aletaha D, Alasti F and Smolen JS. Rheumatoid arthritis near remission: clinical rather than laboratory inflammation is associated with radiographic progression. *Ann Rheum Dis* 2011; 70: 1975–1980.
- Landewe RB, Strand V, Conaghan PG, *et al.* Damage and progression on radiographs in individual joints: data from pivotal randomized controlled trials. *J Rheumatol* 2011; 38: 2018–2022.
- Van der Heijde D and Boyesen P. Measuring disease activity and damage in arthritis. In: Biljsma JWJ and Hachulla E (eds) *EULAR textbook on rheumatic diseases*. London: BMJ, 2012, pp.1107–1134.
- Aletaha D and Smolen JS. The definition and measurement of disease modification in inflammatory rheumatic diseases. *Rheum Dis Clin North Am* 2006; 32: 9–44, vii.
- Scott D, Van Riel PC, van der Heijde DM, *et al.* The EULAR Standing Committee for International Clinical Studies Including Therapeutic Trials. *Assessing disease activity in rheumatoid arthritis: the EULAR handbook of standard methods*. Zürich: EULAR, 1993.
- Bauer EM, Ben-Artzi A, Duffy EL, *et al.* Joint-specific assessment of swelling and power Doppler in obese rheumatoid arthritis patients. *BMC Musculoskelet Disord* 2017; 18: 99.
- Fuchs HA and Pincus T. Reduced joint counts in controlled clinical trials in rheumatoid arthritis. *Arthritis Rheum* 1994; 37: 470–475.
- Aletaha D, Neogi T, Silman AJ, *et al.* 2010 rheumatoid arthritis classification criteria: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2010; 69: 1580–1588.
- Austin PC and Steyerberg EW. The number of subjects per variable required in linear regression analyses. *J Clin Epidemiol* 2015; 68: 627–636.
- Saleem B, Brown AK, Keen H, *et al.* Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. *Ann Rheum Dis* 2011; 70: 792–798.
- Smolen JS, Breedveld FC, Eberl G, *et al.* Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum* 1995; 38: 38–43.
- Backhaus M, Burmester GR, Gerber T, *et al.*; Working Group for Musculoskeletal Ultrasound in the EULAR Standing Committee on International Clinical Studies including Therapeutic Trials. Guidelines for musculoskeletal ultrasound in rheumatology. *Ann Rheum Dis* 2001; 60: 641–649.
- Wakefield RJ, Balint PV, Szudlarek M, *et al.* Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005; 32: 2485–2487.
- Sunk IG, Amoyo-Minar L, Stamm T, *et al.* Interphalangeal osteoarthritis radiographic simplified (iOARS) score: a radiographic method to detect osteoarthritis of the interphalangeal finger joints based on its histopathological alterations. *Ann Rheum Dis* 2014; 73: 1983–1989.
- Gärtner M, Mandl P, Radner H, *et al.* Sonographic joint assessment in rheumatoid arthritis: associations with clinical joint assessment during a state of remission. *Arthritis Rheum* 2013; 65: 2005–2014.

17. Pincus T. Limitations of a quantitative swollen and tender joint count to assess and monitor patients with rheumatoid arthritis. *Bull NYU Hosp Jt Dis* 2008; 66: 216–223.
18. Salaffi F, Filippucci E, Carotti M, *et al.* Inter-observer agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey scale ultrasonography—a preliminary study. *Rheumatology (Oxford)* 2008; 47: 54–58.
19. Cheung PP, Gossec L, Mak A, *et al.* Reliability of joint count assessment in rheumatoid arthritis: a systematic literature review. *Semin Arthritis Rheum* 2014; 43: 721–729.
20. Foltz V, Gandjbakhch F, Etchepare F, *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. Witt M, Mueller F, Nigg A, *et al.* Relevance of grade 1 grayscale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. *Arthritis Rheum* 2013; 65: 1694–1701.
22. Mandl P, Balint PV, Brault Y, *et al.* Metrologic properties of ultrasound versus clinical evaluation of synovitis in rheumatoid arthritis: results of a multicenter, randomized study. *Arthritis Rheum* 2012; 64: 1272–1282.
23. Mandl P, Naredo E, Wakefield RJ, *et al.*; OMERACT Ultrasound Task Force. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *J Rheumatol* 2011; 38: 2055–2062.
24. Terslev L, Naredo E, Aegerter P, *et al.* Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce-part 2: reliability and application to multiple joints of a standardised consensus-based scoring system. *RMD Open* 2017; 3: e000427.
25. Padovano I, Costantino F, Breban M, *et al.* Prevalence of ultrasound synovial inflammatory findings in healthy subjects. *Ann Rheum Dis* 2016; 75: 1819–1823.
26. Millot F, Clavel G, Etchepare F, *et al.* Musculoskeletal ultrasonography in healthy subjects and ultrasound criteria for early arthritis (the ESPOIR cohort). *J Rheumatol* 2011; 38: 613–620.
27. Mandl P, Balint PV, Brault Y, *et al.* Clinical and ultrasound-based composite disease activity indices in rheumatoid arthritis: results from a multicenter, randomized study. *Arthritis Care Res (Hoboken)* 2013; 65: 879–887.

Visit SAGE journals online
[journals.sagepub.com/
 home/tab](http://journals.sagepub.com/home/tab)

 SAGE journals