Original Article

Concordance between ultrasound joint synovitis and clinical joint assessments by patients or physicians in rheumatoid arthritis

Running title: Concordance between US and clinical joint assessment

Ayako Hirata, MD, Takehisa Ogura, MD, Norihide Hayashi, MD, Sayaka Takenaka, MD, Hideki Ito, MD, Kennosuke Mizushina, MD, Yuki Fujisawa, MD, Naoko Yamashita, MD, Sumie Nakahashi, MD, Munetsugu Imamura, MD, Rie Kujime, MD, and Hideto Kameda, MD,PhD

Toho University, Tokyo, Japan

Correspondence author:

Hideto Kameda, Division of Rheumatology, Department of Internal Medicine, Toho University Ohashi Medical Center, 2-17-6 Ohashi, Meguro-ku, Tokyo 153-8515, Japan

Tel.: +81-3-3468-1251, Fax: +81-3-3468-5082

E-mail: hideto.kameda@med.toho-u.ac.jp

Funding: This work was partly supported by the Japan Society for the Promotion of Science KAKENHI (Outline of Grants-in-Aid for Scientific Research), grant no. 15K09535.

Conflict of interest statement: The authors declare that they have no competing interests.

Word count: Abstract 250, manuscript 1669

Tables: 3, Figures: 0, References 30

Abstract

Objective: Ultrasonography (US) has been prevalently used as a valid and objective modality for joint examination in patients with rheumatoid arthritis (RA). This study aimed to examine and compare the concordance between joint symptom, tenderness, or swelling and US synovitis.

Methods: Fifty patients with RA (84% female; median age, 69 years; disease duration, 2.4 years; disease activity score of 28 joints, 3.84) completed the self-evaluation of joint symptoms including pain and considerable stiffness in the (proximal) interphalangeal, metacarpophalangeal, wrist, elbow, shoulder, knee, and ankle joints. These joints were also subjected to physical examination by a physician to evaluate for the presence of tenderness and/or swelling, and to US examination for the presence of synovitis defined as gray-scale score ≥2 or power Doppler signal score ≥1.

Results: In a total of 1492 evaluated joints, symptoms, tenderness, and swelling were observed in 288 (19.3%), 182 (12.2%), and 220 (14.7%) joints, respectively, and US synovitis was observed in 317 (21.2%) joints. The overall concordance rate with US synovitis was the least for joint tenderness ($\kappa = 0.30$) when compared with joint symptoms ($\kappa = 0.39$) or swelling ($\kappa = 0.43$), irrespective of the evaluated joint except for the elbow. Moreover, the percentage of inflamed joints clinically detected only on

the basis of symptom, tenderness, or swelling in patients positive for US synovitis was 4.0%, 0.5%, and 1.8%, respectively.

Conclusion: Joint examination for the presence of swelling is crucial, and patient-reported joint symptoms may be a better clinical assessment than the examination for tenderness.

Significance and Innovations

- Joint examination for the presence of swelling is crucial, and the concordance between joint tenderness and US synovitis was poor.
- Patient-reported joint symptoms may be a better clinical assessment than the examination for tenderness.

Introduction

The integrated measurement of disease activity and the treat-to-target strategy has successfully improved the outcomes of patients with rheumatoid arthritis (RA) [1]. Although the measurement of RA disease activity by using validated composite measures such as DAS (disease activity score) [2], DAS28 (disease activity score of 28 joints) [3], and SDAI (simplified disease activity index) [4] has been well concordant with joint destruction at a patient group level, discordances were frequently observed at the individual patient level [5]. Incorporation of patient-reported outcome, all-or-none assessment of joint examination, and the influence of anemia and hypergammaglobulinemia on the erythrocyte sedimentation rate (ESR), for example, have been proposed in the development or modification of other composite activity measures [6,7].

Ultrasound (US) and magnetic resonance imaging (MRI) are superior to clinical examination in the detection of joint inflammation [8], and US synovitis has been reported to predict the radiographic progression in RA patients in clinical remission [9]. Several studies have already reported on the agreement between US and physical examination in detecting synovial inflammation in patients with RA, indicating about 10% residual US synovitis in clinically nonarthritic joints [10,11]. However, the

importance of the self-assessment of each joint by RA patients in comparison with the evaluator's joint examination for the concordance with US synovitis has not been elucidated.

Therefore, we investigated which component of joint examination, either by patients or by evaluators, is more or less concordant with US synovitis to improve the joint evaluation of patients with RA in daily clinical practice.

Patients and methods

Fifty patients with RA were enrolled in this retrospective study. The following patients were included: (i) patients who fulfilled the American College of Rheumatology (ACR) 1987 revised criteria [12] or the ACR/European League Against Rheumatism (EULAR) 2010 classification criteria [13]; (ii) patients visiting the Division of Rheumatology, Toho University Ohashi Medical Center, from January 2012 to January 2015; (iii) patients who completed the self-assessment of joint pain and considerable joint stiffness for a total of at least 30 joints (those evaluated with DAS28 and ankle joints); and (iv) patients who completed the US joint examinations of at least 28 joints evaluated with DAS28. Other medical information obtained from the records included age, sex, disease

duration from the RA diagnosis, patient's global assessment of pain and disease activity as measured by using a 100-mm visual analogue scale, Health Assessment Questionnaire disability index, the presence of tenderness in 30 joints, the presence of swelling in 30 joints, ESR, serum C-reactive protein level, the presence of rheumatoid factor and anti-cyclic citrullinated peptide antibodies (anti-CCP), and the treatment for RA around the time of the US examination. This study was approved by the institutional ethical committee (approval no. 15-14), and informed consent was obtained from all participants.

US examination

A Xario (Toshiba Medical Systems, Tochigi, Japan) US machine equipped with a multifrequency linear array probe (7–14 MHz) was used. The power Doppler (PD) settings were as follows: PD pulse repetition frequency 16.5 kHz (flow range, 3.8 cm/s); Doppler frequency 6.1 MHz, and low wall filter. Color gain was set just below the level at which noise appeared. The US examination was performed according to the EULAR guidelines for musculoskeletal US in rheumatology [14], by one of three rheumatologists and ultrasonographers certified by the Japan College of Rheumatology (AH, TO, or NH). These examiners evaluated at least 30 of the joints described above

by using longitudinal and transverse scans, without knowledge of the other medical information of the patients.

The US examiners performed the final scoring of the recorded US findings in a blinded manner according to the OMERACT definitions [15]. The gray scale (GS) was graded semiquantitatively on a scale of 0–3 (0 = absent, 1 = mild, 2 = moderate, and 3 = marked) in a combined measure of synovial hypertrophy and fluid retention of the articular recess [16–18]. The intra-articular PD signals were also graded on a scale of 0–3. Joints or tendons graded as GS ≥ 2 or PD ≥ 1 were judged as having joint synovitis [19,20]. The interobserver reliability of the US scoring was examined by using the recorded US findings of six randomly selected patients, resulting in Kendall's W coefficient of concordance of 0.838.

Statistical analysis

Statistical analysis was performed by using EZR software (version 1.25; Saitama Medical Center, Jichi Medical University, Saitama, Japan) [21], which is a graphical user interface for R (version 3.1.1; The R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were summarized with medians and ranges, and analyzed by using the Mann–Whitney U test, whereas binominal data from the three

groups were examined by using chi-square test. The concordance rates and unweighted kappa (κ) values were calculated separately between US examination, patients' self-evaluation of the joints, and clinical joint examination. The rate of concordance was analyzed by using Cohen's κ statistic. The κ coefficients were divided as follows: <0.0 = poor, 0–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial, and 0.81–1.0 = almost perfect agreement [11]. P values <0.05 were considered statistically significant.

Results

Patient characteristics

A total of 50 RA patients were enrolled in this study. Their median age was 69 years, and 84% of them were female (Table 1). The median disease duration was 31 months. The median tender joint count (TJC) of 30 joints, swollen joint count (SJC) of 30 joints, and DAS28 was 2, 3, and 3.8, respectively. Anti-CCP antibody was positive in 78% of the patients, and 95% of them showed a high (more than three times than the upper limit of normal) positivity. Among the patients, 54% were receiving methotrexate and 18% were receiving biological agents.

Prevalence of clinical and US joint findings and the concordance between clinical and US findings

The US finding of 30 joints, including the left and right ankle joints, were available in 46 (92%) of the 50 enrolled patients. In a total of 1492 evaluated joints, symptoms, tenderness, and swelling were observed in 288 (19.3%), 182 (12.2%), and 220 (14.7%) joints, respectively, and US synovitis was observed in 317 (21.2%) joints. In general, the concordance between the presence of US synovitis and the clinical findings in each joint were slight or fair (Table 2). Swelling showed the best concordance with US synovitis, whereas tenderness showed the worst concordance in all the examined joints except for the elbow joint. Interestingly, the patients' self-assessed joint symptoms showed a slightly better concordance than did swelling in the metacarpophalangeal (κ = 0.42 and κ = 0.36, respectively) and wrist (κ = 0.50 and κ = 0.42, respectively) joints.

US synovitis clinically detected only with self-assessed joint symptoms, and physician-assessed tenderness and swelling

Finally, we compared the exclusive detection rate of US synovitis by each clinical joint assessment to determine which assessment could be omitted in daily clinical practice (Table 3). Surprisingly, the detection rate according to joint symptom alone was better

than that according to swelling alone, and tenderness again showed the worst detection rate (4.0%, 1.8%, and 0.5%, respectively; $P < 0.01$).

Discussion

In our study, swelling showed the best concordance with US synovitis, whereas tenderness showed the worst concordance in all the examined joint regions except for the elbow. Moreover, self-reported joint symptoms showed an exclusive significance in detecting 4% of US synovitis; however, tenderness was useful only in 0.5% of the cases.

The concordance between the presence of tenderness or swelling and US synovitis in this study was similar to those of previous reports with comparable patient populations in terms of sample size and disease activity [10,11,22,23]. The discrepancy between clinical examinations and US synovitis may be due to the presence of osteoarthritis and/or joint deformity, anatomical complexity, and the cutoff level of each assessment. US examination has been reported to improve patient self-assessment of the joints [24] and even physician's assessment of synovitis [25]. Therefore, US examination in daily clinical practice may be useful in improving clinical joint assessments, both at present and in the future.

Validated composite measures such as DAS, DAS28, and SDAI include patient-reported outcomes, evaluator's assessments, and acute phase reactants [2–4]. However, pain is the most important determinant of RA disease activity according to patients' perception, whereas it is mostly joint swelling for the evaluation by physicians. Indeed, the elimination of TJC and patient global assessment from composite activity measures improved the correlation with MRI synovitis and more accurately predicted radiographic progression [6]. However, because of the importance of the patients' perspective in the clinical assessment of RA and other rheumatic diseases [26], as well as the acceptable agreement between the joint assessment by patients and physicians [27], we examined and compared the relative importance of the patients' self-assessment of each joint and the physicians' assessment of joint tenderness in this study. Our results demonstrated the superiority of patient self-assessment over the evaluation for tenderness when US synovitis was considered a gold standard. These findings have been further supported by the very low rate of exclusive detection of US synovitis on the basis of joint tenderness (0.5%) as compared with patients' self-assessment of the joints (4.0%). Considering the time required for the assessment of joint tenderness, we conclude that this parameter can be excluded from the clinical joint assessments, or may be substituted by patient-reported joint symptoms in daily clinical

practice. Another approach to improve the clinical assessment of RA disease activity is by including US joint assessment in composite measures such as DAS28 [28]; however, this is time consuming and unlikely to be acceptable in daily clinical practice.

The limitations of this study include its small sample size, retrospective analysis of the data, and the absence of other high-sensitivity imaging such as MRI [29,30]. Nonetheless, our results suggested the inclusion of patients' self-assessment of their joints instead of the joint examination for the presence of tenderness in the composite activity measures of RA.

Acknowledgment

This work was partly supported by the Japan Society for the Promotion of Science KAKENHI (Outline of Grants-in-Aid for Scientific Research), grant no. 15K09535.

References

1. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2016;75:3–15. doi:10.1136/annrheumdis-2015-207524.

2. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916–20.

3. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL, et al. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.

4. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology, 2003;42:244–57.

5. Kameda H, Kanbe K, Sato E, Ueki Y, Saito K, Nagaoka S, et al. A merged presentation of clinical and radiographic data using probability plots in a clinical trial, the JESMR study. Ann Rheum Dis 2013;72:310–2.

6. Baker JF, Conaghan PG, Smolen JS, Aletaha D, Shults J, Emery P, et al. Development and validation of modified disease activity scores in rheumatoid arthritis: superior correlation with magnetic resonance imaging-detected synovitis and radiographic progression. Arthritis Rheum 2014;66:794–802.

7. Kameda H, Sekiguchi N, Nagasawa H, Amano K, Takei H, Suzuki K, et al. Development and validation of the handy rheumatoid activity score with 38 joints (HRAS38) for rheumatoid arthritis in patients receiving infliximab. Mod Rheumatol 2006;16:381–8.

8. Colebatch AN, Edwards CJ, Østergaard M, van der Heijde D, Balint PV, D'Agostino MA, et al. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. Ann Rheum Dis 2013;72:804–14.

9. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum 2008;58:2958–67.

10. Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K, et al. Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. Arthritis Rheum.2013;65:890–8.

11. Ceponis A, Onishi M, Bluestein HG, Kalunian K, Townsend J, Kavanaugh A. Utility of the ultrasound examination of the hand and wrist joints in the management of established rheumatoid arthritis. Arthritis Car Res 2014;66:236–44.

12. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

13. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69:1580–8.

14. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;60:641–9.

15. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol. 2005;32:2485–7.

16. Naredo E, D'Agostino MA, Wakefield RJ, Moller I, Balint PV, Filippucci E, et al. Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. Ann Rheum Dis. 2013;72:1328–34.

17. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum. 2003;48:955–62.

18. Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2005;52:733–43.

19. Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K, et al. Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League against rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. Arthritis Rheum. 2013;65:890–8.

20. Witt M, Mueller F, Nigg A, Reindl C, Leipe J, Proft F, et al. Relevance of grade 1 gray-scale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. Arthritis Rheum 2013;65:1694–701.

21. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant. 2013;48:452–8.

22. Garrigues F, Jousse-Joulin S, Bouttier R, Nonent M, Bressollette L, Saraux A. Concordance between clinical and ultrasound findings in rheumatoid arthritis. J Bone Spine 2013;80:597–603.

23. Cheung PP, Ruyssen-Witrand A, Gossec L, Paternotte S, Le Bourlout C, Mazieres M, et al. Reliability of patient self-evaluation of swollen and tender joints in rheumatoid arthritis: a comparison study with ultrasonography, physician, and nurse assessments. Arthritis Care Res 2010;62:1112–9.

24. Cheung PP, Lahiri M, Teng GG, Lim AY, Lau TC, Lateef A, et al. A randomized controlled trial for improving patient self-assessment of synovitis in rheumatoid arthritis with education by ultrasonography: the RAEUS Study. Rheumatology 2015;54:1161–9.

25. Ogasawara M, Murakami G, Yamada Y, Nemoto T, Kageyama M, et al. Autofeedback from ultrasound images provides rapid improvement in palpation skills for identifying joint swelling in rheumatoid arthritis. J Rheumatol 2012;39:1207–14.

26. Van Tuyl LH, Boers M. Patient-reported outcomes in core domain sets for rheumatic diseases. Nat Rev Rheumatol 2015;11:705–12.

27. Janta I, Naredo E, Martínez-Estupiñán L, Nieto JC, De la Torre I, Valor L, et al. Patient self-assessment and physician's assessment of rheumatoid arthritis activity: which is more realistic in remission status? A comparison with ultrasonography. Rheumatology 2013;52:2243–50.

28. Damjanov N, Radunovíc G, Prodanovíc S, Vukovíc V, Milić V, Pasalic KS, et al. Construct validity and reliability of ultrasound disease activity score in assessing joint inflammation in

RA: comparison with DAS-28. Rheumatology 2012;51:120–8.

29. Bøyesen P, Haavardsholm E, van der Heijde D, Østergaard M, Hammer HB, Sesseng S, et al. Prediction of MRI erosive progression: a comparison of modern imaging modalities in early rheumatoid arthritis patients. Ann Rheum Dis 2011;70:176–9.

30. Tamai M, Kawakami A, Iwamoto N, Kawashiri SY, Fujikawa K, Aramaki T, et al. Comparative study of the detection of joint injury in early-stage rheumatoid arthritis by magnetic resonance imaging of the wrist and finger joints and physical examination. Arthritis Care Res 2011;63:436–9.

Table 1. Patient characteristics

Age (years)	69 (27–90)
Sex (female:male)	n = 42:n = 8
RA duration (months)	31 (0–608)
TJC of 30 joints	2 (0–23)
SJC of 30 joints	3 (0–20)
DAS28	3.8 (0.8–8.3)
HAQ-DI	0.4 (0–2.8)
Anti-CCP positive	n = 39 (78%)
MTX use	n = 27 (54%)
Biologics use	n = 9 (18%)

Values are median (range) unless indicated otherwise.

RA, rheumatoid arthritis; TJC, tender joint count; SJC, swollen joint count; DAS28, disease activity score of 28 joints; HAQ-DI, Health Assessment Questionnaire disability index; CCP, cyclic citrullinated peptide; MTX, methotrexate.

Table 2. Concordance (κ value) between US synovitis and clinical joint findings

		IP/PIP (n = 500)	MCP (n = 500)	Wrist (n = 100)	Elbow (n = 100)	Shoulder (n = 100)	Knee (n = 100)	Ankle (n =92)	Total (n = 1492)
Symptoms	%	16.8	19.2	38.0	27.0	11.0	20.0	12.8	19.3
	κ vs. US	0.27	0.42	0.50	0.21	0.35	0.46	0.39	0.39
Tenderness	%	11.8	10.4	25.0	9.0	15.0	13.0	9.6	12.2
	κ vs. US	0.26	0.26	0.21	0.45	0.34	0.39	0.21	0.30
Swelling	%	10.4	13.2	42.0	9.0	12.0	18.0	22.3	14.7
	κ vs. US	0.36	0.36	0.42	0.52	0.44	0.45	0.47	0.43
US synovitis	%	11.2	24.4	51.0	15.0	25.0	31.0	18.1	21.2

Values are % of joints positive for symptoms, tenderness, swelling and US synovitis, respectively, and κ values versus US synovitis.

US, ultrasound; IP/PIP, interphalangeal/proximal interphalangeal; MCP, metacarpophalangeal joint.

Table 3. US-positive joints clinically detected only according to symptoms, tenderness, and swelling

	IP/PIP (n = 500)	MCP (n = 500)	Wrist (n = 100)	Elbow (n = 100)	Shoulder (n = 100)	Knee (n = 100)	Ankle (n = 92)	Total (n = 1492)
Sy(+) T(-) S(-)	12 (2.4%)	25 (5.0%)	9 (9.0%)	3 (3.0%)	5 (5.0%)	4 (4.0%)	1 (1.0%)	59 (4.0%)
Sy(-) T(+) S(-)	2 (0.4%)	3 (0.6%)	1 (1.0%)	0	0	1 (1.0%)	0	7 (0.5%)
Sy(-) T(-) S(+)	5 (1.0%)	11 (2.2%)	6 (6.0%)	0	0	2 (2.0%)	3 (3.0%)	27 (1.8%)

Values are number (%) of joints.

US, ultrasound; IP/PIP, interphalangeal/proximal interphalangeal; MCP, metacarpophalangeal joint; Sy, symptoms; T, tenderness; S, swelling.