Article:
Takase-Minegishi, K, Horita, N, Kobayashi, K et al. (2018) Diagnostic test accuracy of ultrasonography for synovitis in rheumatoid arthritis: systematic review and meta-analysis. Rheumatology, 57 (1). pp. 49-58. ISSN: 1462-0324

https://doi.org/10.1093/rheumatology/kex036

Diagnostic test accuracy of ultrasonography for synovitis in rheumatoid arthritis: systematic review and meta-analysis

Kaoru Takase-Minegishi[1], Nobuyuki Horita[2], Kouji Kobayashi[1], Ryusuke Yoshimi[3], Yohei Kirino[3], Shigeru Ohno[1], Takeshi Kaneko[2], Hideaki Nakajima[3], Richard J. Wakefield[4], and Paul Emery[4]

1 Center for Rheumatic Diseases, Yokohama City University Medical Center, Yokohama, Japan

2 Department of Pulmonology, Yokohama City University Graduate School of Medicine, Yokohama, Japan

3 Department of Stem Cell and Immune Regulation, Yokohama City University Graduate School of Medicine, Yokohama, Japan

4 Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Correspondence to Nobuyuki Horita, MD, PhD

3-9 Fukuura, Kanazawa-Ku, Yokohama, 236-0004, Japan.

Tel: +81-45-352-7962, Fax: +81-45-352-7963

E-mail: horitano@yokohama-cu.ac.jp

Keywords: rheumatoid arthritis, ultrasonography, magnetic resonance imaging

Abstract

Objective. To evaluate diagnostic test accuracy of ultrasonography (US) compared with magnetic resonance imaging (MRI) for the detection of synovitis in rheumatoid arthritis (RA) patients.

Methods. A systematic literature search was performed in the Pubmed, EMBASE, the Cochrane Library, and Web of Science Core Collection. Studies evaluating diagnostic test accuracy of US for synovitis detected by MRI as the reference standard for wrist, metacarpophalangeal (MCP), proximal interphalangeal (PIP), and knee joints were included. To assess the overall accuracy, we calculated the diagnostic odds ratio (DOR) using a DerSimonian-Laird random-model and area under the hierarchical summary receiver operating characteristics (AUC) using Holling's proportional hazard models. The summary estimate of the sensitivity and the specificity were obtained using the bivariate model.

Results. Fourteen of 601 identified articles were included in the review. The DOR was 11.6 (95%CI 5.6-24, $I^2 = 0\%$), 28 (95%CI 12-66, $I^2 = 11\%$), 23 (95%CI 6.5-84, $I^2 = 19\%$), 5.3 (95%CI 0.60-48, $I^2 = 0\%$) and AUC was 0.81, 0.91, 0.91, 0.61, for wrist, MCP, PIP, and knee joints, respectively. The summary estimate of sensitivity and specificity were 0.73 (95%CI 0.51-0.87)/0.78 (95%CI 0.46-0.94), 0.64 (95%CI 0.43-0.81)/0.93 (95%CI 0.88-0.97), 0.71 (95%CI 0.33-0.93)/0.94 (95%CI 0.89-0.97), and 0.91 (95%CI 0.56-0.99)/0.60 (95%CI 0.20-0.90) for wrist, MCP, PIP, and knee joints, respectively.

Conclusion. US is a valid and reproducible technique for detecting synovitis in the wrist and finger joints. It may be considered for routine use as part of the standard diagnostic tool in RA.

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by autoimmunity and polyarticular synovial inflammation; it subsequently causes bone destruction. For patients with RA the current concept is "treat-to-target" with clinical remission the primary treatment goal aiming to achieve it as soon as possible [1]. Clinical trials have demonstrated that early treatment reduces inflammation, resulting in limited structural change and better long-term outcomes [2-6]. Therefore, early diagnosis of RA is essential for initiation of treatment. Recently, advances in the field of imaging techniques have resulted in ultrasonography (US) and magnetic resonance imaging (MRI) being recommended for making the diagnosis and monitoring the disease activity in RA patients [7]. US and MRI have been shown to be more sensitive than clinical examination in detecting synovitis, both in active disease and in remission. [8-10]. The predictive value of evaluating subclinical synovitis by imaging techniques was first described by Brown et al [11] and it has been demonstrated that US-detected subclinical synovitis can lead to radiographic progression, even in clinical remission [12]. Moreover, the presence of inflammation observed with US or MRI can be used to predict the progression from undifferentiated inflammatory arthritis to clinical RA [13-16].

Although MRI is capable of directly visualising joint inflammation, there are difficulties in performing MRI as an initial test because of the limited resources. The assessment of multiple joints with MRI is time-consuming and expensive for routine use. By contrast, US is relatively low-cost, non-invasive, and has real-

time capabilities and portability. Despite these advantages, there are some limitations of this technology, whilst several studies have highlighted the ability of US in the detection of joint inflammation as compared with MRI, there was considerable discrepancy of results in these previous studies, and US is considered to be an operator-dependent technology. To assist in resolving this discrepancy, this systematic review and meta-analysis was conducted.

Methods

Overview

The study protocol followed the Cochrane Handbook for Diagnostic Test Accuracy Review and the Preferred Reporting Items in Systematic Reviews, and the Meta-analysis statement has been registered on the international prospective register of systematic reviews (PROSPERO) as number 42016033912 [17-20]. Institutional Review Board approval and patient informed consent were waived due to review nature of this study.

Both case-control and cohort studies were included when they provided sufficient data for both sensitivity and specificity of US for detection of MRI-judged synovitis in human RA. However, no eligible case-control study were found. Here, single- and two-gate studies were customarily termed cohort and case-control studies. Studies covering only sensitivity or only specificity were excluded. Non-English written reports and conference abstracts were allowed in the protocol, though none of them were eventually eligible.

Search strategy

In the electronic search, we systematically searched Pubmed, EMBASE, the Cochrane Central Register of Controlled Trials, and Web of Science Core Collection. Search formulas were presented as supplementary data (Supplementary Text 1). References of previously published reviews and those of included original studies were checked through the hand search. Two investigators (KM, NH)

independently screened the candidate articles by checking the title and abstract after uploading the citation list into the software, Endnote X7 (Thomson Reuters, Philadelphia, USA). After independent screening, articles still regarded as candidates by at least one investigator were then scrutinised independently through full-text reading. Final inclusion were decided after resolving discrepancies between the two investigators.

Participants

We included patients with the diagnosis of RA defined by the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria or 1987 ACR criteria for RA [21, 22]. Synovitis in RA at wrist joints, metacarpophalangeal (MCP) joints, proximal interphalangeal (PIP) joints, and knee joints were the target pathology. Neither bone erosion nor synovitis that was caused by connective tissue diseases other than RA was included in this study.

Index and reference test

The index test was US in any mode including color Doppler US, power Doppler US, B-mode US, grey scale US, 2D US, 3D US, and contrast-enhanced US [23]. Positive and negative results for US were determined based on judgement by the authors of the original researches. When a report presented diagnostic test accuracy of two US modes separately, we used the only data of PD to avoid duplicate use of the data from the same subject. In such a case, we selected PD rather than GS because recent data suggested

that PD can provide more accurate data than GS for synovitis in RA [7].

Reference test were MRI in any mode including non-enhanced MRI, enhanced MRI, dynamic MRI, 1.5-Tesla MRI, and 3-Tesla MRI, compact MRI, low-field extremity MRI, 0.2-Tesla MRI [24]. Positive and negative result in MRI were also determined based on judgement by the authors of the original researches. We categorised the quality of MRI based on MRI mode as follows: high = high field contrast-enhanced MRI, moderate = high agreement was confirmed in comparison with high field contrast-enhanced MRI, low = low field extremity MRI. Four cohorts, using MRI without contrast enhancement, evaluated the ability to detect synovitis in comparison with conventional 1.5 T contrast-enhanced MRI in RA patients in a preliminary study [33, 35, 41]. Therefore, we defined “moderate quality”.

Primary outcome

Co-primary outcomes were diagnostic test accuracy of US for synovitis diagnosed by MRI using following statistics: diagnostic odds ratio (DOR), area under hierarchical summary receiver operating characteristic (HSROC) curve (AUC), the summary estimates of the sensitivity, the specificity, the positive likelihood ratio (PLR), and the negative likelihood ratio (NLR). Wrist, MCP, PIP, and knee joints were evaluated separately [17, 18].

Risk of bias

The two investigators independently evaluated each study by scoring seven domains of A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) evaluation sheet [25]. Any discrepancies were resolved through discussion.

Data synthesis

Data were crosschecked after extracted by the two investigators independently. Then, we composed a two by two contingency.

All analyses were done based on numbers of joints but not on numbers of patients.

We used both the HSROC model and bivariate model. To determine the overall diagnostic test accuracy, we calculated the DOR using the DerSimonian-Laird random-effect model and the AUC using bivariate model of Reitsma [26, 27]. Heterogeneity was indicated by I^2 wherein 0% means no heterogeneity and 100% means the strongest heterogeneity. We obtained a paired forest plot, HSROC curve, and summary estimates of the sensitivity and the specificity using the bivariate model. PLR and NLR were obtained using the summary estimate of the sensitivity and the specificity [26, 27]. DOR, AUC, and HSROC were obtained from all the cohorts regardless of the cut-off value. Summary estimates of the sensitivity, summary estimate of specificity, PLR, and NLR were obtained from cohorts that used US cutoff values between negative and positive. According to the authors, five adaptive cutoff scores of US were used as follows: (1) 0 grouped as "negative", 1 grouped as "positive"; (2) 0 grouped as "negative", 1-2 grouped as "positive"; (3) 0 grouped

as “negative”, 1-3 grouped as “positive”; (4) 0-1 grouped as “negative”, 2-3 grouped as “positive”; (5) 0-1 grouped as “negative”, 2-4 grouped as “positive”.

We conducted subgroup analyses based on US modes and MRI modes.

We used the following commands in the "mada" package of the statistics software R: the “madauni” command for DOR and the “reitsma” command for the AUC, the HSROC curve, the summary estimates of sensitivity and specificity [26, 27]. Review Manager 5.3 (Cochrane, London, UK) was used to draw the paired forest plot and the Cochrane risk of bias graph.

Interpretation of diagnostic test accuracy statics

AUC was interpreted in a four-grade scale as follows: AUC < 0.75 not accurate, 0.75 < AUC < 0.92 good, 0.93 < AUC < 0.96 very good, AUC < 0.97 excellent [28]. PLR value in the range of < 2, between 2 and 5, between 5 and 10, and > 10 were recognised as a not meaningful, small, moderate, and large increase in probability [29]. NLR in the range of > 0.5, between 0.2 and 0.5, between 0.1 and 0.2, and < 0.1 were interpreted as a not meaningful, small, moderate, and large decrease of probability [29].

Results

Study search and study characteristics

Of the 601 candidate articles, we finally identified 14 eligible reports [30-43]. Three of them presented two cohorts, thus we included 17 independent cohorts (Figure 1). To obtain data that were not presented in each original report, we tried to have a contact with authors of 18 reports. Among them, authors of 3 original reports provided additional information [32-34].

Among the included 14 reports, six were from Japan, four were from Denmark, and one were from each of Belgium, China, Germany, and UK. Publication dates ranged from 2001 to 2014. All reports used one-gate cohort recruiting method (Table 1). One was English written letter and the other were English full articles. Seven, three, and one were conducted in single university hospital, in multi-center hospital-based arthritis clinics, and in single hospital respectively, while three reports did not provide specific information of their facility. To diagnose RA, one used both 1987 ACR criteria and 2010 ACR/EULAR criteria, 12 used 1987 ACR criteria only, and one did not provided the information of diagnostic criteria. Numbers of patients in each study ranged from 6 to 77 with a median of 18, with a total of 376 (Table 1). Concerning Cochrane risk of bias evaluation, one study had high risk of index bias due to arbitrary decision of US cutoff [36]. No other report had any high risk of bias or any high applicability concern (Supplementary Figure 1).

Among 17 cohorts, 12 used non-enhanced power Doppler, three used grey-scale US, one used contract-enhanced power Doppler, one used both grey-scale US and power Doppler. Wrist, MCP, PIP, and

knee joint were evaluated in five, 12, six, and two cohorts, respectively and were evaluated for 275, 2060, 1073, and 31 joints, respectively (Table 2). Median sensitivities/specificities were 0.66/0.90, 0.77/0.96, 0.80/0.91, and 0.77/0.55 for wrist, MCP, PIP, and knee joints, respectively (Figure 2).

Wrist

Five cohorts with 275 wrist joints yielded DOR of 11.6 (95%CI 5.6-24, I^2 = 0%) and AUC of 0.81. This AUC suggested that US had good diagnostic test accuracy for wrist synovitis (Figure 3, Table 2).

Using the cutoff value between absence and presence, the summary estimate of sensitivity and specificity were 0.73 (95%CI 0.51-0.87) and 0.78 (95%CI 0.46-0.94), respectively. Based on PLR of 3.3 and NPV of 0.35, both positive and negative US results suggested a small change of synovitis probability (Table 2 A).

MCP

Data of 2060 MCP joints from 12 cohorts suggested DOR of 28 (95%CI 12-66, I^2 = 11%) and AUC of 0.91, which meant that US had good diagnostic test accuracy for MCP synovitis (Figure 3, Table 2).

When applying the cutoff value between absence/presence, the summary estimate of sensitivity and specificity were 0.64 (95%CI 0.43-0.81) and 0.93 (95%CI 0.88-0.97), respectively (Table 2). PLR was 9.1 (95%CI 4.2-19) suggesting moderate increase of MCP synovitis probability when US detected it.

PIP

Six cohorts of 1073 PIP joints yielded a DOR of 23 (95%CI 6.5-84, I^2 = 19%) and AUC of 0.91. This AUC value suggested that US had good diagnostic test accuracy for PIP synovitis (Figure 3, Table 2).

Using the data from five cohorts that used a cut-off value between absence/presence, the summary estimate of sensitivity and specificity were 0.71 (95%CI 0.33-0.93) and 0.94 (95%CI 0.89-0.97), respectively. Positive and negative US results suggested large and small change of synovitis probability, respectively (Table 2).

Knee

The diagnostic test accuracy of knee was researched in a smaller number of cohorts and joints compared to other joints. DOR was 5.3 (95%CI 0.60-48, I^2 = 0%) and AUC was 0.61, which indicated that the US did not have good diagnostic test accuracy for knee synovitis (Figure 3, Table 2). The 95% CI of both PLR and NLR included 1.0, which meant no diagnostic value. (Table 2).

MRI mode subgroup analysis

We carried out subgroup analyses focusing on studies with high-quality MRIs and those with moderate- or high-quality MRIs. These analyses almost replicated the results from studies with any MRI

modes (Supplementary Table 1).

US mode subgroup analysis

Based on US mode subgroup analysis, power Doppler US showed better overall diagnostic testaccuracy than grey-scale US (Supplementary Table 2). Notably, power Doppler US had very good AUC to detect MRI proven synovitis in MCP and PIP joints. Power Doppler US positive with a cutoff value between absence/presence or 0/1 largely increase the probability of MRI proven synovitis in MCP and PIP joints (Supplementary Table 2).

Discussion

US is widely used for the evaluation of RA inflammatory activity in daily practice and in clinical trials. Despite the increasing availability of US application, there remains a lack of quality validation studies. The Outcome Measures in Rheumatology (OMERACT) group has proposed definitions for synovial fluid and synovial hypertrophy [44]. US allows visualisation of the pannus developing in the inflamed joint. Grey-scale and Doppler US are capable of measuring synovial proliferation and vascularity, respectively. Several approaches for assessing synovitis in RA patients have been described in published studies. Qualitative, semiquantitative and scoring systems have been used for assessing synovitis by grey-scale and/or Doppler US.

Our systematic review and meta-analysis provided the evidence supporting the use of US for evaluating synovitis in RA patients. MRI mode based subgroup analyses suggested the robustness of our analysis. We showed that the diagnostic test accuracy of US was good for detecting synovitis at joint level using MRI as the reference standard, especially with regard to MCP and PIP joints. The data suggest that US of wrist joints was less accurate than MCP and PIP joints. The diagnostic test accuracy for knee joints was low, but was based on a small number of cohorts. Although it has limited resolution for deeper joints and the patient's body habitus may sometimes make examination difficult, US has been shown to be more sensitive than clinical examination in determining synovitis for large joints such as the shoulder and knee [45, 46]. The small sample size increased the size of confidence intervals, and therefore a greater statistical

uncertainty of the results, even when the diagnostic test has a high sensitivity.

This meta-analysis has several limitations. The number of papers qualifying for the analysis is low and we used data from direct communication with the original authors. Therefore, recall bias would occur. Our systematic review focused on wrist, finger and knee joints. As noted above, only two reports representing three cohorts compared the ability of US and MRI to detect synovitis for knee joints. However, the small joints of the hands and feet play a central role in the diagnosis of RA. Our systematic review shows that US can be recommended as a reliable diagnostic tool for synovitis in RA. Previous systematic review suggested that the wrist, MCP and metatarsophalangeal (MTP) joints should be scanned in the diagnostic process of RA [47]. Despite the fact that feet were not evaluated in this study, similar results may be obtained for MTP joints. As, MRI is not a gold standard to detect synovitis without contrast enhancement, we carried out subgroup analyses based on MRI quality. MRI is also reader-dependent particularly when an established scoring method such as RAMRIS is not used. Furthermore, subgroup analysis based on US mode showed Doppler US had very good diagnostic test accuracy and was more accurate than grey-scale US concerning detecting MCP synovitis. Some subgroup analyses provided imprecise estimation of test accuracy due to limited number of studies.

Our systematic review did not distinguish early and established RA. In this meta-analysis, all of identified eligible studies were performed for established RA patients. Harman et al. assessed the efficacy of US compared with contrast-enhanced MR in patients with newly diagnosed RA [48]. However, as they

showed only sensitivity and specificity data, this study was excluded. Another issue is operator-dependent techniques for scoring systems. Although US examination for synovitis is mostly carried out from the dorsal aspect of the finger joint, several studies have addressed volar synovitis. Moreover, our systematic review revealed a lack of consensus regarding standardised US scoring system for synovitis. The definition of a "positive" or "negative" US-determined synovitis was defined with different cut-off values. The sensitivity and specificity of a quantitative test are dependent on the cut-off value above or below, and there is a "trade-off" between sensitivity and specificity. We chose the bivariate model to determine the overall diagnostic test accuracy of US. This model takes into account the potential trade-off between sensitivity and specificity by explicitly incorporating this negative correlation in the analysis, with the result it could calculate the DOR/AUC. However, the reliability of the estimated accuracy is limited, especially for knee, where only a limited number of studies is being present. In addition, the optimal cut-off value was not determined in this study. Although five adaptive cutoff scores were used, it was not enough to make the distinction at various cut-off scores due to the small sample size.

In summary, this systematic review and meta-analysis suggest that US, especially power Doppler US, is a valid and reproducible technique for detecting synovitis in the wrist and finger joints. US has certain great advantages over MRI, including low cost, portability, and lack of contraindications. It requires consideration of appropriate training and quality assessment. However, US may allow more widespread, therefore be considered for routine use as part of the standard diagnostic tool in RA.

Acknowledgements

The authors would like to thank Erika Ota, St.Luke's international university Graduate school of nursing science, for help with the literature search. The authors declare no conflict of interest regarding this work. Dr. Takase-Minegishi is supported by grants from Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (Grant No. 15K19578). Dr. Yoshimi is supported by grants from Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (Grant No. 26461468) and Yokohama Foundation for Advancement of Medical Science. Dr. Kirino is supported by grants from Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (Grant No. 26713036 and 15K15374).

Rheumatology key messages

This is the first systematic review and meta-analysis on US assessment of synovitis in RA patients.

US seems a valid and reproducible technique for detecting synovitis in the wrist and finger joints.

Further US quality assessment is necessary for diagnostic test accuracy.

References

1. Smolen JS, Breedveld FC, Burmester GR, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis. 2016;75:3-15.

2. Lard LR, Visser H, Speyer I, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. Am J Med. 2001;111:446-51.

3. Nell VP, Machold KP, Eberl G, et al. Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. Rheumatology (Oxford). 2004;43:906-14.

4. Schipper LG, Vermeer M, Kuper HH, et al. A tight control treatment strategy aiming for remission in early rheumatoid arthritis is more effective than usual care treatment in daily clinical practice: a study of two cohorts in the Dutch Rheumatoid Arthritis Monitoring registry. Ann Rheum Dis 2012;71:845-50.

5. Rantalaiho V, Kautiainen H, Korpela M, Targeted treatment with a combination of traditional DMARDs produces excellent clinical and radiographic long-term outcomes in early rheumatoid arthritis regardless of initial infliximab. The 5-year follow-up results of a randomised clinical trial, the NEO-RACo trial. Ann Rheum Dis. 2014;73:1954-61.

6. Markusse IM, Akdemir G, Dirven L, et al. Long-Term Outcomes of Patients With Recent-Onset

Rheumatoid Arthritis After 10 Years of Tight Controlled Treatment: A Randomized Trial. Ann Intern Med. 2016;164:523-31.

7. D'Agostino MA, Terslev L, Wakefield R, et al. Novel algorithms for the pragmatic use of ultrasound in the management of patients with rheumatoid arthritis: from diagnosis to remission. Ann Rheum Dis. 2016 Aug 23. [Epub ahead of print]

8. Szkudlarek M, Narvestad E, Klarlund M, et al. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. Arthritis Rheum. 2004;50:2103-12.

9. Scheel AK, Hermann KG, Ohrndorf S, et al. Prospective 7 year follow up imaging study comparing radiography, ultrasonography, and magnetic resonance imaging in rheumatoid arthritis finger joints. Ann Rheum Dis. 2006;65:595-600.

10. Brown AK, Quinn MA, Karim Z, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum. 2006;54:3761-73.

11. 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.

12. Yoshimi R, Hama M, Takase K, et al. Ultrasonography is a potent tool for the prediction of progressive

joint destruction during clinical remission of rheumatoid arthritis. Mod Rheumatol. 2013;23:456-65.

13. Eshed I, Feist E, Althoff CE, et al. Tenosynovitis of the flexor tendons of the hand detected by MRI: an early indicator of rheumatoid arthritis. Rheumatology (Oxford) 2009;48:887-91.

14. Salaffi F, Ciapetti A, Gasparini S, et al. A clinical prediction rule combining routine assessment and power Doppler ultrasonography for predicting progression to rheumatoid arthritis from early-onset undifferentiated arthritis. Clin Exp Rheumatol 2010;28:686-94.

15. Filer A, De Pablo P, Allen G, et al. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. Ann Rheum Dis 2011;70:500-7.

16. Machado PMMC, Koevoets R, Bombardier C, et al. The value of magnetic resonance imaging and ultrasound in undifferentiated arthritis: a systematic review. J Rheumatol 2011;87:31-7.

17. Leeflang MM, Deeks JJ, Takwoingi Y, et al. Cochrane diagnostic test accuracy reviews. Syst Rev 2013;2:82.

18. Petra M, Constantine G, Jonathan D, Roger H, Yemisi T. Cochrane Handbook for Diagnostic Test Accuracy Reviews: Chapter 10 Analysing and Presenting Results (ver 1.0). 2010. Available at http://dta.cochrane.org/handbook-dta-reviews. Accessed on 10th August, 2016.

19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

20. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev 2012;1:2.

21. 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;69:1580-8.

22. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum1988;31:315-24.

23. Kang T, Lanni S, Nam J, et al. The evolution of ultrasound in rheumatology. Ther Adv Musculoskelet Dis 2012;4:399-411.

24. Borrero CG, Mountz JM, Mountz JD. Emerging MRI methods in rheumatoid arthritis. Nat Rev Rheumatol 2011;7:85-95.

25. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155:529-536.

26. Doebler P. Package 'mada'. Available from: https://cran.r-project.org/web/packages/mada/mada.pdf#search='mada+r'. Accessed on November 5th, 2015.

27. Doebler P., & Holling H. Meta-analysis of diagnostic accuracy with mada. Available from: https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf#search='mada+r+systematic+review+diagnostic

+accuracy'. Accessed on November 5th, 2015.

28. Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg 2005;9:16-20.

29. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet 2005;365:1500-1505.

30. Beckers C, Jeukens X, Ribbens C, et al. (18)F-FDG PET imaging of rheumatoid knee synovitis correlates with dynamic magnetic resonance and sonographic assessments as well as with the serum level of metalloproteinase-3. Eur J Nucl Med Mol Imaging 2006;33:275-80.

31. Freeston JE, Brown AK, Hensor EM, et al. Extremity magnetic resonance imaging assessment of synovitis (without contrast) in rheumatoid arthritis may be less accurate than power Doppler ultrasound. Ann Rheum Dis 2008;67:1351.

32. Fukuba E, Yoshizako T, Kitagaki H, et al. Power Doppler ultrasonography for assessment of rheumatoid synovitis: comparison with dynamic magnetic resonance imaging. Clin Imaging 2013;37:134-7.

33. Horikoshi M, Suzuki T, Sugihara M, et al. Comparison of low-field dedicated extremity magnetic resonance imaging with articular ultrasonography in patients with rheumatoid arthritis. Mod Rheumatol 2010;20:556-60.

34. Kamishima T, Tanimura K, Shimizu M, et al. Monitoring anti-interleukin 6 receptor antibody treatment for rheumatoid arthritis by quantitative magnetic resonance imaging of the hand and power Doppler

ultrasonography of the finger. Skeletal Radiol 2011;40:745-55.

35. Ogishima H, Tsuboi H, Umeda N, et al. Analysis of subclinical synovitis detected by ultrasonography and low-field magnetic resonance imaging in patients with rheumatoid arthritis. Mod Rheumatol 2014;24:60-8.

36. Scheel AK, Hermann KG, Kahler E, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum 2005;52:733-43.

37. Szkudlarek M, Court-Payen M, Strandberg C, et al. Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. Arthritis Rheum 2001;44:2018-23.

38. Szkudlarek M, Court-Payen M, Strandberg C, et al. Contrast-enhanced power Doppler ultrasonography of the metacarpophalangeal joints in rheumatoid arthritis. Eur Radiol 2003;13:163-8.

39. Szkudlarek M, Klarlund M, Narvestad E, et al. Ultrasonography of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis: a comparison with magnetic resonance imaging, conventional radiography and clinical examination. Arthritis Res Ther 2006;8:R52.

40. Takase K, Ohno S, Takeno M, et al. Simultaneous evaluation of long-lasting knee synovitis in patients undergoing arthroplasty by power Doppler ultrasonography and contrast-enhanced MRI in comparison with histopathology. Clin Exp Rheumatol 2012;30:85-92.

41. Taniguchi D, Tokunaga D, Oda R, et al. Maximum intensity projection with magnetic resonance

imaging for evaluating synovitis of the hand in rheumatoid arthritis: comparison with clinical and ultrasound findings. Clin Rheumatol 2014;33:911-7.

42. Terslev L, Torp-Pedersen S, Savnik A, et al. Doppler ultrasound and magnetic resonance imaging of synovial inflammation of the hand in rheumatoid arthritis: a comparative study. Arthritis Rheum 2003;48:2434-41.

43. Xiao H, Liu M, Tan L, et al. Value of ultrasonography for diagnosis of synovitis associated with rheumatoid arthritis. Int J Rheum Dis 2014;17:767-75.

44. Wakefield RJ, Balint P, Szkudlarek M et al.: Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2485-7.

45. Naredo E, Aguado P, De Miguel E, et al. Painful shoulder: comparison of physical examination and ultrasonographic findings. Ann Rheum Dis 2002;61:132-136.

46. Kane D, Balint PV, Sturrock RD. Ultrasonography is superior to clinical examination in the detection and localization of knee joint effusion in rheumatoid arthritis. J Rheumatol. 2003;30:966-71.

47. Ten Cate DF, Luime JJ, Swen N, et al. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis--a systematic review of the literature. Arthritis Res Ther 2013;15:R4.

48. Harman H, Tekeoğlu İ, Sağ MS, et al. Diagnostic value of musculoskeletal ultrasound in newly diagnosed rheumatoid arthritis patients. Turk J Phys Med Rehab 2015;61:326-32.

Figure Legends

Figure 1. PRISMA flow diagram.

Figure 2. Paired forest plots.

MCP: metacarpophalangeal. PIP: proximal interphalangeal. TP: true positive. FP: false positive. FN: false negative. TN: true negative.

Figure 3. Hierarchical summary receiver operating characteristic curves.

MCP: metacarpophalangeal. PIP: proximal interphalangeal.