Article:
Tan, YK, Allen, JC, Lye, WK et al. (3 more authors) (2017) Dichotomous versus semi-quantitative scoring of ultrasound joint inflammation in rheumatoid arthritis using novel individualized joint selection methods. Clinical Rheumatology, 36 (5). pp. 1137-1141. ISSN 0770-3198

https://doi.org/10.1007/s10067-016-3432-5

TITLE PAGE

Title: Dichotomous versus semi-quantitative scoring of ultrasound joint inflammation in rheumatoid arthritis using novel individualized joint selection methods

Authors:

York Kiat Tan, John C. Allen Jr, Weng Kit Lye, Philip G Conaghan, Li-Ching Chew, Julian Thumboo

ABSTRACT

Objective

The aim of the study is to compare the responsiveness of two joint inflammation scoring systems (dichotomous scoring (DS) versus semi-quantitative scoring (SQS)) using novel individualized ultrasound joint selection methods and existing ultrasound joint selection methods.

Methods

Responsiveness measured by the standardized response means (SRMs) using the DS and the SQS system (for both the novel and existing ultrasound joint selection methods) were derived using the baseline and the 3 months total inflammatory scores from 20 rheumatoid arthritis patients. The relative SRM gain ratios (SRM-Gains) for both scoring system (DS and SQS) comparing the novel to the existing methods were computed.

Results

Both scoring systems (DS and SQS) demonstrated substantial SRM-Gains (ranged from 3.31 to 5.67 for the DS system and ranged from 1.82 to 3.26 for the SQS system). The SRMs using the novel methods ranged from 0.94 to 1.36 for the DS system and ranged from 0.89 to 1.11 for the SQS system. The SRMs using the existing methods ranged from 0.24 to 0.32 for the DS system and ranged from 0.34 to 0.49 for the SQS system.

Conclusions

The DS system appears to achieve high responsiveness comparable to SQS for the novel individualized ultrasound joint selection methods.

Key Indexing Term: Rheumatoid Arthritis, Ultrasonography, Synovitis, Joints

AFFILIATION(S) AND ADDRESS(ES) OF THE AUTHOR(S):

1) Y.K. Tan, MBBS (Singapore), MMED (Singapore), MRCP (UK). Consultant, Department of Rheumatology and Immunology, Singapore General Hospital, Singapore. Adjunct Assistant Professor, Duke-NUS Graduate Medical School, Singapore. Clinical Senior Lecturer, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

2) J.C. Allen Jr, PhD. Office of Clinical Sciences, Centre for Quantitative Medicine, Duke-NUS Graduate Medical School Singapore.

3) W.K. Lye, MSc. Office of Clinical Sciences, Centre for Quantitative Medicine, Duke-NUS Graduate Medical School Singapore.

4) P.G. Conaghan. MBBS, PhD, FRACP, FRCP. Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, & NIHR Leeds Musculoskeletal Biomedical Research Unit, UK.

5) L.C. Chew. B. MED. SCI (UK), B.M.B.S. (UK), M.R.C.P. (London, UK), F.A.M.S. (Rheumatology, Singapore), F.R.C.P. (London, UK). Senior Consultant, Department of Rheumatology and Immunology, Singapore General Hospital, Singapore. Adjunct Assistant Professor, Duke-NUS Graduate Medical School, Singapore. Clinical Senior Lecturer, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

6) J. Thumboo. MBBS (Singapore), MMED (Singapore), MRCP (UK), FAMS (Rheumatology, Singapore), FRCP (Edin). Senior Consultant, Department of Rheumatology and Immunology, Singapore General Hospital, Singapore. Professor, Duke-NUS Graduate Medical School, Singapore. Adjunct Professor, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

Correspondence:

York Kiat Tan. Department of Rheumatology and Immunology, Singapore General Hospital, Outram Road, Singapore 169608. Tel: +65 63214028. Email: tan.york.kiat@singhealth.com.sg

INTRODUCTION

Musculoskeletal ultrasound is well suited as an outcome measure tool for rheumatoid arthritis (RA) joint assessment [1]. As demonstrated by two prospective multi-center RA studies [2,3], ultrasound synovitis dichotomous scoring (DS) and semi-quantitative scoring (SQS) are at least as valid and reliable as clinical examination based on the Outcome Measures in Rheumatology (OMERACT) filter. SQS requires grading across multiple severity catergories which may lead to greater scoring variability than DS (yes/no grading), DS has the advantage of being relatively quicker to perform. The acceptable performance metrics of these scores may depend on their application, either in clinical trials or everyday clinical care. It is therefore important to examine the relative performance metrics of both scoring methods, using a range of ultrasound scoring systems (that typically evaluate different joints). Recently, SQS assessment of joint synovitis, in combination with novel ultrasound joint selection methods, demonstrated superior sensitivity for detecting response to treatment when compared to existing methods. The improvement in sensitivity may be attributed to selection of a greater proportion of affected joints for subsequent scanning, using the novel methods, which maximized the potential for detecting improvement resulting in greater sensitivity for detecting change [4]. In this pilot study, our objective was to compare sensitivity of the DS and SQS joint inflammation scoring systems using novel and existing ultrasound joint selection methods.

MATERIALS AND METHODS

In this IRB approved study, seropositive RA subjects (DAS28 >3.2, ≥ 5 tender and/or swollen joints) initiating or up-titrating their disease modifying anti-rheumatic drugs (DMARDs) and corticosteroid treatments were recruited and followed up prospectively. At baseline and 3 months, 44-joint clinical and ultrasound assessment data (on bilateral shoulders, hips, elbows, knees, ankles, mid-foot, wrists and small joints of the hand and feet as previously described [4]), were collected for analysis. Patients were recruited from March 2013 to April 2015 after providing informed consent.

Ultrasound and clinical joint assessment

Ultrasonography was performed using Philips Medical Systems EPIQ 5G machine with a multi-frequency

linear array transducer (5–17 MHz) or General Electric Healthcare LOGIQe machine with a multi-frequency linear array transducer (5–13 MHz).The ultrasound settings were pre-set for the joints for each machine and the same machine (and probe) was used for each patient throughout the study period. The ultrasound scans were acquired and scored by a rheumatologist (blinded to the joint assessor's clinical findings) experienced in musculoskeletal ultrasound. Clinical joint assessments were performed by either a study nurse or a metrologist. Standardized scanning based on the EULAR guidelines [5] was performed in the same outpatient location. Power Doppler (PD) vascularity and grey-scale synovial hypertrophy (GSSH) at the joints were scored using SQS or DS systems. The SQS system scores these inflammatory changes on a 0-3 severity scale using methods described previously [4] while the DS system scores the changes as Yes (equivalent to SQS score of 1-3) or No (equivalent to SQS score of 0). Ultrasound tenosynovitis, clinical joint tenderness and swelling were scored dichotomously in both systems. Joint pain was scored dichotomously with DS system or semi-quantitatively (0-3 severity scale) with SQS system.

Novel and existing methods

The novel and existing joint selection methods were previously described [4] and are summarized below. The novel methods include (i) the Individualized Ultrasound method, which selects up to 7 or 12 ultrasonographically most inflamed joints, and (ii) the Individualized Composite-Ultrasound method, which utilizes both clinical joint symptoms and ultrasound joint findings to select up to 7 or 12 target joints. The existing methods include (a) a pre-specified 7-joint count [6] and (b) a 12-joint count derived from an ultrasound data reduction method [7].

For Individualized Ultrasound and existing methods, the individual joint score (IJS) at each joint site was calculated as the sum of component ultrasound sub-scores divided by the maximum possible score at the joint. For Individualized Composite-Ultrasound method, the IJS at each joint site was calculated as the sum of component ultrasound and clinical sub-scores divided by the maximum possible score at the joint, The IJSs were computed so as to equalize the score weights across the joints [4].

“The joint selection process using the novel methods follows an algorithm whereby the IJSs from 44 joint sites (including bilaterally the metacarpophalangeal joints 1 to 5, first interphalangeal joints, proximal interphalangeal joints 1 to 5, metatarsophalangeal joints 1 to 5, wrists, ankles, mid-tarsals, elbows, knees, shoulders and hips) are ranked from the largest to the smallest. Joints with a larger IJS are selected first and the selection proceeds through a pre-determined joint sequence (e.g. from smaller joint sites then followed by medium joint sites and finally larger joint sites) [4]. Eventually, joints with lesser IJS will be selected until the target joint count (7 or 12) is reached. Therefore, the novel scoring methods enable selection of up to 7 or 12 most affected joint sites for subsequent scanning. The relevant IJSs from these joint sites are then further computed and analyzed as described in the section on statistical analysis below. “

Statistical analysis:

For the novel and existing methods, total inflammatory score (TIS) per patient was calculated as the sum of the IJSs at the relevant joint sites. Responsiveness at 3 months was measured using the standardized response mean (SRM). Mean TIS at 0 and 3 months were obtained and used to derive the SRM which was computed as the mean change in the TIS score divided by the standard deviation (SD) of the change in the TIS score. Threshold values from Cohen for effect size (ES) are often used for interpreting the SRM [8-10] i.e. ES <0.20 is trivial; 0.20≤ES<0.50 is small; 0.50≤ES<0.80 is moderate; and ES≥0.80 is large. Relative gains in SRM magnitude (SRM-Gains) at 3 months were reported comparing novel to existing methods using the 7 and 12 joints approaches (e.g. for both DS and SQS systems, the SRM-Gain for Individualized Composite-Ultrasound method using the 7 joints approach was calculated as the ratio of Individualized Composite-Ultrasound method’s SRM and existing method’s SRM for the 7 joints approach).

RESULTS

Patient baseline characteristics

Twenty RA subjects (mean age 57.2 years; 90% female; 75% Chinese; mean (SD) disease duration 49.9 (60.4) months; baseline mean (SD) ESR 40.8 (28.3) mm/hr) had mean DAS28 scores at baseline and 3

months of 4.76 and 4.19, respectively.

DS and SQS systems

The mean (SD) total inflammatory scores of the DS and SQS systems (at baseline and 3 months) using the novel and existing methods are shown in figure 1 (via the 7 joints approach) and figure 2 (via the 12 joints approach).The SRMs for both scoring systems were similar in magnitude for both Individualized Composite-Ultrasound and Individualized Ultrasound methods, for both the 7 and 12 joints approaches.

For the 7 joints approach, both scoring systems when used with the novel methods (Individualized Composite-Ultrasound and Individualized Ultrasound) showed higher SRMs (ranged from 0.94 to 1.36) when compared to existing methods (SRMs ranged from 0.24 to 0.34). The SRM-Gains ranged from 3.26 to 5.67.

For the 12 joints approach, both scoring systems when used with the novel methods (Individualized Composite-Ultrasound and Individualized Ultrasound) also showed higher SRMs (ranged from 0.89 to 1.16) when compared to the SRMs of existing methods (ranged from 0.32 to 0.49). The SRM-Gains ranged from 1.82 to 3.63. (Table 1)

DISCUSSION

Our study is the first showing DS system can be highly sensitive to change, with a degree of responsiveness comparable to SQS system, using novel joint selection methods. In contrast, both scoring systems had modest sensitivity to change using existing methods.

Three studies have compared DS and SQS quantifying ultrasound synovitis in RA patients [2,3,6]. DS versus SQS was compared in 68 RA patients using 20, 28 and 38 pre-fixed joints [2]. Intra-rater reliability was comparable between DS and SQS (with moderate to good intra-class correlation coefficient (ICC) ranging from 0.53 to 0.97). Another study involving 62 RA subjects also reported comparable ICC between DS and SQS (mean ICC for GS and PD ultrasound was 0.85 and 0.80, respectively) [3]. A third study, however, demonstrated higher intra-rater reliability with DS (mean kappa 0.83) when compared to

SQS (mean kappa 0.64) using four images with synovitis and erosions [6]. The inter-rater reliability was also higher for DS (kappa of 0.62 and 0.84 for GS and PD synovitis, respectively) when compared to SQS (kappa of 0.55 and 0.67 for GS and PD synovitis, respectively) using thirty-three ultrasound images. DS of ultrasound joint inflammation was found to correlate well with MRI findings in two studies [11,12] demonstrating evidence of construct validity. In the first study involving 22 RA subjects, ultrasound joint inflammation scored dichotomously on the feet [11] had higher percentage exact agreement with MRI findings than clinical examination (63.6-77.3% versus 45.5–59.1%, respectively). In the second study involving 12 RA subjects, ultrasound detected joint effusion and synovitis scored dichotomously at the finger joints had sensitivity and specificity of 0.83 and 0.94, respectively, when compared to MRI [12].

DS may have lesser variability than SQS as demonstrated in one study showing higher intra/inter-rater reliability [6] for SQS. Additionally, DS may be logistically easier to implement in RA studies than SQS (e.g. easier to train personnel(s) in its use). However, there is less information on DS relative to SQS. Conceivably, DS may require a longer time interval than SQS before improvement in joint inflammation is detected.

Limitations of our study are the small sample size in this pilot study, the absence of reliability testing of DS versus SQS (although these were examined by previous studies) and evaluation of change using two time-points. Future larger scale studies could assess change over multiple time-points.

Overall, our study demonstrated high sensitivity to improvement using the DS system, which was shown to be comparable in sensitivity to the SQS system when used in combination with the novel individualized joint selection methods. If confirmed in larger RA cohorts, this would have important implications when choosing an ultrasound scoring system in future RA studies.

ACKNOWLEDGEMENTS

The authors thank colleagues and staff from the study site for the support and help provided in this study. This study was funded by SingHealth Duke-NUS Academic Medical Centre, through the Medicine

Academic Clinical Programme Grant: Nurturing Clinician Scientist Scheme. We would also like to acknowledge the “Estate of Tan Sri Khoo Teck Puat” for the funding support (through the Khoo Clinical Scholars Pilot Award) and the National University of Singapore.

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflicts of interest

ETHICAL STANDARD STATEMENT

This study has been approved by the appropriate ethics committee and all subjects provided their informed consent prior to their inclusion in the study. Details that might disclose the identity of the subjects under study are omitted.

REFERENCES

1. Tan YK, Conaghan PG. Imaging in rheumatoid arthritis. Best Pract Res Clin Rheumatol 2011;25:569-84
2. Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J, et al. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. Ann Rheum Dis 2010;69:828-33
3. Mandl P, Balint PV, Brault Y, Backhaus M, D'Agostino MA, Grassi W, et al. Metrologic properties of ultrasound versus clinical evaluation of synovitis in rheumatoid arthritis: results of a multicenter, randomized study. Arthritis Rheum 2012;64:1272-82
4. Tan YK, Allen JC Jr, Lye WK, Conaghan PG, D'Agostino MA, Chew LC, et al. Novel Ultrasound Joint Selection Methods Using a Reduced Joint Number Demonstrate Inflammatory Improvement when Compared to Existing Methods and Disease Activity Score at 28 Joints. J Rheumatol 2016;43:34-7
5. 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
6. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum 2009;61:1194-201
7. Naredo E, Rodríguez M, Campos C, Rodríguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve- joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. Arthritis Rheum 2008;59:515-22
8. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178–89
9. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care 2002;2:e15
10. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239–46

11. Wakefield RJ, Freeston JE, O'Connor P, Reay N, Budgen A, Hensor EM, et al. The optimal assessment of the rheumatoid arthritis hindfoot: a comparative study of clinical examination, ultrasound and high field MRI. Ann Rheum Dis 2008;67:1678-82
12. Rahmani M, Chegini H, Najafizadeh SR, Azimi M, Habibollahi P, Shakiba M. Detection of bone erosion in early rheumatoid arthritis: ultrasonography and conventional radiography versus non-contrast magnetic resonance imaging. Clin Rheumatol 2010;29:883-91