Construct validity and response to therapy of the U9 ultrasonographic scale for assessment of disease activity in rheumatoid arthritis

Mohamed Mortada[1] ID , Hany Aly[2] ID , Reem Elmallah[3] ID , Ahmed Radwan[4] ID , Ahmed Elsaman[4] ID

[1]Department of Rheumatology and Rehabilitation, Zagazig University, Egypt
[2]Department of Rheumatology, Al-Hussein Hospital, Al-Azhar University, Cairo, Egypt
[3]Department of Rheumatology and Rehabilitation, Ain Shams University, Cairo, Egypt.
[4]Department of Rheumatology and Rehabilitation, Sohag University, Egypt

Abstract

Objectives: To test the construct validity of the U9 ultrasonographic scale, to determine the cut-off points for different degrees of rheumatoid arthritis (RA) activity, and to determine whether or not US assessment with the U9 score is useful for monitoring the response to treatment of RA.

Material and methods: A prospective, multicenter study was conducted in 4 different centers in Egypt. All RA patients who were recruited were subject to evaluation of clinical disease activity by the Clinical Disease Activity Index (CDAI) and Disease Activity Score of 28 joints based on erythrocyte sedimentation rate (DAS28-ESR). Assessment of the Functional Status by the Health Assessment Questionnaire (HAQ) and U9 ultrasound score was performed. All the targeted joints were assessed by EULAR recommendations and the combined score of EULAR/OMERACT (0–3). Targeted tendons scored 0–3. After three months of treatment, CDAI and DAS28-ESR, HAQ, and U9 were repeated to detect the response.

Results: One hundred and forty patients with mean age 39.26 ±11.30 were recruited from 4 centers. With regard to convergent validity, the U9 ultrasonographic scale was significantly associated with clinical parameters (CDAI and DAS28-ESR) as well as functional state (HAQ) at both visits. Likewise, concerning discriminative validity, the U9 scale showed the ability to distinguish different grades of RA activity, presenting well-defined cut-off points of different grades (severe, moderate, and mild), with very good specificity and sensitivity (11.5, 5.5, and 3.5, respectively). A significant parallel decrease was detected in clinical and sonographic scales at the follow-up assessment.

Conclusions: The U9 ultrasound scale showed good construct (convergent and discriminative) validity and can be used to monitor the disease and therapeutic response to treatment in RA.

Key words: rheumatoid arthritis, ultrasound, score.

Introduction

Evaluation of rheumatoid arthritis (RA) disorder activity is of primary importance with significant implications for clinical decisions [1]. Correct analysis of disease activity of RA remains a challenging process.

Several clinical scores and indices have been used to assess disease activity in both clinical practice and trials: the Disease Activity Score based on 28 joints (DAS28) [2], Clinical Disease Activity Index (CDAI) [3], Simplified Disease Activity Index (SDAI) [4], American College of Rheumatology (ACR) response criteria [5], and Boolean-based remission criteria [6].

Most of the previously mentioned indices showed great reliability [7]; nonetheless, many common items are

subjective such as global patient assessment, physician global assessment, and number of tender joints.

Musculoskeletal ultrasonography (US) indicates direct objective signs of inflammation of the synovial lining of joints as well as surrounding tendons and soft tissues. Several studies have demonstrated that ultrasound is more sensitive than clinical examination in assessing disease activity in rheumatoid patients [8].

In order to reduce examination time, many scores have been proposed, selecting different fixed sets of joints. There are many proposed sets of composite ultrasonographic scores [9]; however, a fixed set of joints may not be an ideal tool to assess a disease such as RA, which affects many joints and tendons in different presentations.

In a previous study [9], three different composite scores were proposed. The first score (modified U8 score) included the bilateral wrists, the 2nd metacarpophalangeal (MCP), the 3rd MCP and knees. This is the same set of joints proposed by Yoshimi et al. in 2015 [10], with scores for each joint adjusted according to the EULAR/OMERACT combined score, and thus the score range is 0–24.

The second score (U9) was the same as the modified U8 score, in addition to the scoring of the joint or tendon most clinically affected (i.e. single joint or tendon); consequently the score range is 0–27.

The proposed third score (8 + 2) was the same as the modified U8 score plus the scoring of the two most clinically affected joints or tendons (i.e. one joint and one tendon, two joints or two tendons); therefore the score range is 0–30.

All target joints were evaluated by grey-scale (GS) and power Doppler (PD) ultrasound, using the EULAR/OMERACT combined score (0–3). Targeted tendons were scored 0–3, with either a GS scale, or a PD, then the highest score was used. The U9 score was proven to be the most closely correlated with disease activity parameters [11].

The primary objective of the present study was to test the construct validity of the U9 ultrasonographic scale and to determine the cut-off points for different grades of RA activity. The secondary objective was to determine whether or not US assessment with the U9 score is useful for monitoring the response to RA treatment.

Material and methods

A prospective multicenter observational study was conducted from July 2019 to December 2019. Four tertiary referral university hospitals participated in the study. Approvals were obtained from the institutional review board of the four universities, and all participants signed informed consent. The IRB number of the study was #S20-150.

Patients

Patients had to fulfill ACR/EULAR 2010 RA criteria to apply for this study [11]. All recruited patients were subjected to clinical assessment and US examination.

Clinical assessment

RA activity was assessed clinically using the CDAI [3] and DAS28-ESR [2]. The functional status assessment was performed by the Health Assessment Questionnaire (HAQ) [12].

Ultrasound examination

Ultrasonographic assessment using the U9 score includes eight joints (bilateral wrists, 2nd MCP, 3rd MCP and both knees), besides the joint or tendon most clinically affected (joint swelling, tenderness and limitation of range of motion). The method of grading was as follows:
- right 2nd MCP, G (0, 1, 2, 3), left 2nd MCP, G (0, 1, 2, 3),
- right 3rd MCP, G (0, 1, 2, 3), left 3rd MCP, G (0, 1, 2, 3),
- right wrist, G (0, 1, 2, 3), left wrist, G (0, 1, 2, 3),
- right knee, G (0, 1, 2, 3), left knee, G (0, 1, 2, 3).

The joint most clinically affected was selected from among 48 joints (hand proximal interphalangeal (PIPs), MCPs [1, 4, 5], elbows, gleno-humeral joints, acromio-clavicular joints, sterno-clavicular, hips, ankles, metatarsophalangeal joints (MTPs and foot PIPs). Any affected tendons could be selected. Rheumatologists decide to include either one joint or one tendon in the eight fixed joints.

All target joints were evaluated by GS and power Doppler PD ultrasound according to EULAR guidelines [13]. Synovitis was graded in each joint (0–3) using the combined score EULAR/OMERACT [14]. Targeted tendons were scored (0–3) by GS or PD ultrasound according to the US atlas by Hammer et al. [15], and the highest score was used.

Clinical assessments and ultrasound scans were performed in each center by a rheumatologist with 5 to 15 years of experience in musculoskeletal ultrasonography.

All patients received treatment (biologic and non-biologic disease-modifying antirheumatic drugs [DMARDs]), based on the decision of the treating physicians. Physicians were eligible to change/modify treatment according to disease activity. There was no need for specific therapy in the current study.

The disease activity assessment, functional assessment, and ultrasonographic assessment using the U9 score were repeated after three months to detect the response to change after therapy. The sonographers were blinded to clinical data.

Statistical analysis

A statistical data analysis was carried out using version 25 of SPSS (IBM SPSS Statistics), IBM Corporation, Armonk, USA; August 2017. Data were expressed as mean, standard deviation (SD), percentage, and number. As a descriptive value for quantitative results, mean and standard deviation were used.

The paired *t*-test was used to compare the same variable means at various times, and it was also used with the McNemar χ^2 test to compare the same variables' percentages at different times. Two quantitative variables were compared using Pearson's correlation test.

The value of *r* is explained as follows: *r* positive – positive correlation, *r* negative – negative correlation, *r* < 0.4 – weak correlation, 0.4–0.7 – moderate correlation and 0.7–1.0 – strong correlation.

The EULAR/OMERACT ultrasound score's predictive value has been assessed to distinguish between high, moderate, and low-level disease activity and to calculate the most appropriate cut-off levels that provide for maximum accuracy in the receiver operating characteristic curve (ROC curve) (highest sensitivity and specificity at the same time). For all these tests, the level of significance (*p*-value) was adjusted to < 0.05.

Results

Patient characteristics

The baseline demographic data for 140 patients with RA: the mean age was 39.26 ±11.30 years, and the mean disease duration was 23.70 ±14.60 years. Most of the pa-tients were female (*n* = 107, 76.4%; male: *n* = 33, 23.6%). The majority of the study subjects (91.4%) were treated with non-biologic classical disease-modifying antirheumatic drugs (csDMARDs), most of them in combined therapy, *n* = 121 (86.4%), in monotherapy (methotrexate) only 5% (*n* = 7), and 8.6% (*n* = 12) received biological DMARDs.

Convergent validity

The U9 ultrasonographic scale was significantly correlated with clinical parameters (CDAI and DAS28-ESR) and functional state (HAQ) at both visits (Table I).

According to DAS28-ESR, at baseline there were 5 (3.6%) patients in remission, 8 (5.7%) patients had mild activity, 39 (27.9%) patients had moderate activity, and 88 (62.9%) patients had severe activity.

Discriminant validity

The receiver operating characteristic curves (ROC) to discriminate the ability of the score to distinguish between RA patients with severe activity and non-active disease were very good with an AUC of 0.918 (95% CI: 0.885 to 0.951; *p* < 0.001). The receiver operating characteristic curve achieved a maximum Youden's index value of 11.5, where sensitivity was 85.7% and specificity 80.6%.

The receiver operating characteristic curves to discriminate the ability of the score to distinguish between different grades of active RA showed good sensitivity and specificity (Fig. 1). Cut-off values for severe, moderate and mild activities were 11.5, 5.5, and 3.5 respectively (Table II).

Table I. Correlation between clinical and ultrasound findings at the first and second visits

Visit		CDAI	mHAQ	U9 total score
First visit				
DAS28	Pearson's correlation	0.963	0.556	0.806
	p-value	< 0.001	< 0.001	< 0.001
CDAI	Pearson's correlation		0.547	0.787
	p-value		< 0.001	< 0.001
mHAQ	Pearson's correlation			0.431
	p-value			< 0.001
Second visit				
DAS28	Pearson's correlation	0.953	0.395	0.790
	p-value	< 0.001	< 0.001	< 0.001
CDAI	Pearson's correlation		0.317	0.773
	p-value		< 0.001	< 0.001
mHAQ	Pearson's correlation			0.317
	p-value			< 0.001

CDAI – Clinical Disease Activity Index, DAS28 – Disease Activity Score based on 28 joints, mHAQ – multidimensional Health Assessment Questionnaire.
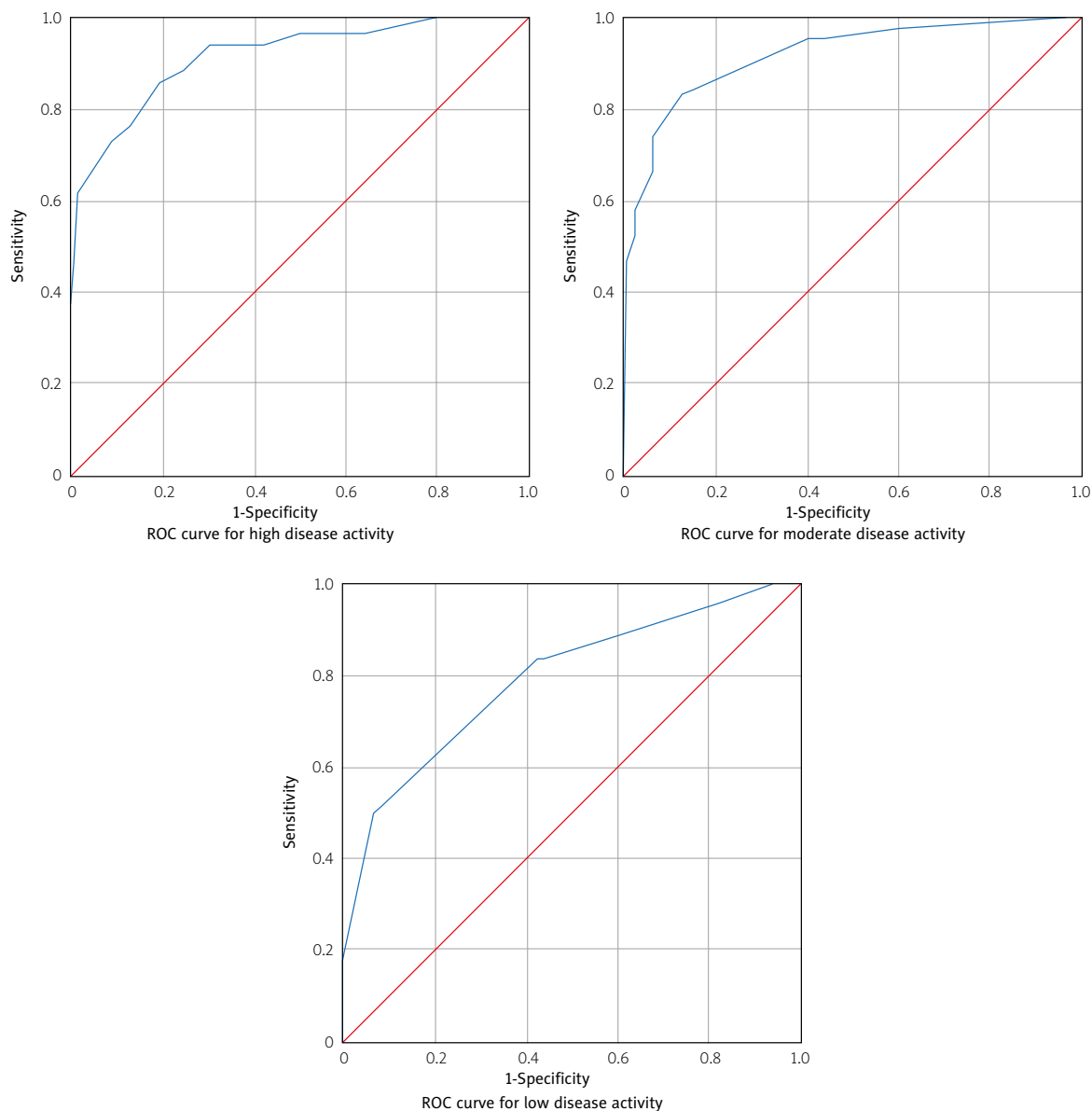
Fig. 1. Receiver operating characteristic (ROC) curve for prediction of disease activity by U9 ultrasound score.

Table II. Cut-off values of U9 ultrasonographic scale for different grades of activity of rheumatoid arthritis

Disease activity	Youden index	Cut-off U9 ultrasonographic scale value	Sensitivity (%)	Specificity (%)
High	0.241	11.5	85.7	80.6
Moderate	0.206	5.5	83.2	88.0
Low	0.460	3.5	83.3	57.1

Comparison between U9 and U8

The receiver operating characteristic curves to compare the ability of U9 and U8 scores (by omitting the most affected joint or tendon) to distinguish between RA patients with different degrees of severity showed that the sensitivity was better for the U9 score compared to the U8 score (85.9% vs. 82.9% for high disease activity cases; 83.2% vs. 80% for moderate disease activity and 83.3% vs. 80.6% for low disease activity, respectively).

Regarding specificity, the U9 showed lower specificity compared to the U8 score for high disease activity cases (80.6% vs. 82.9%; respectively), but was similar regarding moderate and lower disease activity cases, probably indicating that moderate and low disease activity had no or minimal additional symptomatic joints or tendons (Fig. 2 and Table III).

Clinical and sonographic course

Findings on DAS28-ESR, CDAI, and U9 ultrasonographic scale assessed throughout the study are shown in Table IV.

A significant parallel decrease in clinical and sonographic scales was found in the follow-up assessment
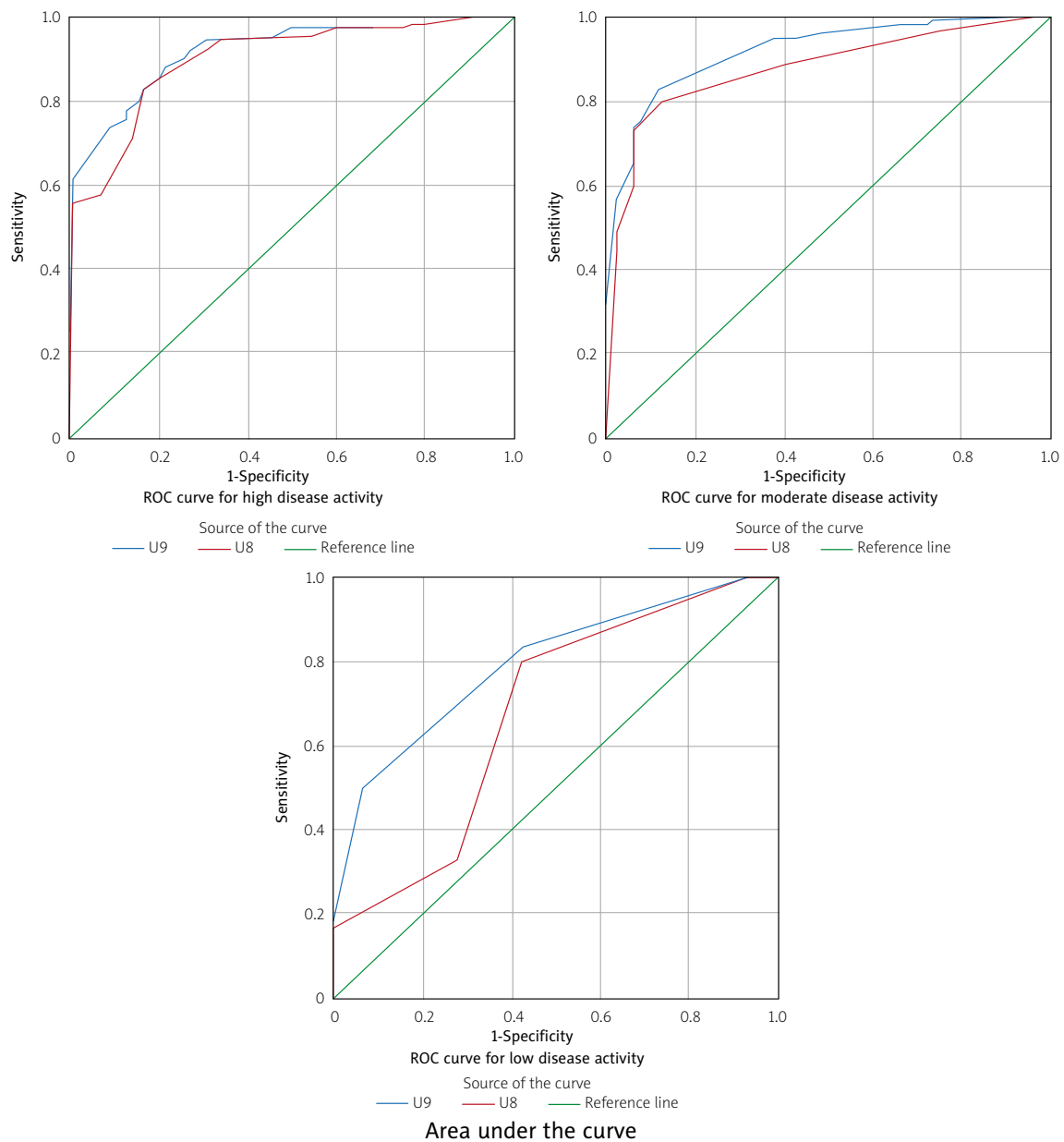

Disease activity		Area	Std. error	p-value	Asymptomatic 95% lower bound	Confidence interval upper bound
High	U9	0.918	0.017	< 0.001	0.885	0.951
	U8	0.897	0.020	< 0.001	0.858	0.935
Moderate	U9	0.928	0.020	< 0.001	0.885	0.963
	U8	0.887	0.025	< 0.001	0.838	0.937
Low	U9	0.790	0.067	0.002	0.658	0.921
	U8	0.676	0.090	0.056	0.498	0.853

Fig. 2. Receiver operating characteristic (ROC) curve analysis for the value of EULAR/OMERACT US score to estimate disease activity.

Table III. Sensitivity statistics of the receiver operating characteristic curve analysis for U9 compared to U8 scores

Disease activity		Youden index	Cut-off OMERACT value	Sensitivity (%)	Specificity (%)
High	U9	0.241	11.5	85.7	80.6
	U8	0.242	10.5	82.9	82.9
Moderate	U9	0.206	5.5	83.2	88.0
	U8	0.233	5.5	80.0	88.0
Low	U9	0.460	3.5	83.3	57.1
	U8	0.471	3.5	80.6	57.1

($p < 0.001$). Of note, the right 5th MCP was the most frequently selected joint to be added in the U9 scale (20.7%) followed by the right ankle (11.7%) and the most frequently selected tendon to be added in U9 was the right extensor carpi ulnaris (8.6%).

Discussion

In the past decade, musculoskeletal US has become a widely used tool in diagnosing and assessing the activity of RA. There are many proposed composite scales to assess RA activity by US [16]. In order to reduce examination time, and to be used in daily practice, a representative fixed set of joints was chosen by the scale. The preliminary data were presented in abstarct poster (AB1119) during the EULAR 2020 Conference [17].

RA is a polyarticular disease that affects many joints and tendons in several presentations. The fixed set of joints may not be appropriate for detecting the exact state of the disease in all affected joints, especially in mild to

Table IV. Mean ±SD values for clinical, laboratory, and power Doppler ultrasonography score parameters at the baseline and follow-up assessments

Item	First visit	Follow-up visit	*p*-value
Clinical assessment			
No tender joints	6.47 ±3.29	3.35 ±2.22	< 0.001
No swollen joints	5.93 ±3.37	2.59 ±1.90	< 0.001
ESR	42.12 ±15.24	26.84 ±12.32	< 0.001
DAS28	5.29 ±1.21	3.95 ±0.99	< 0.001
CDAI	23.00 ±10.15	11.14 ±6.62	< 0.001
mHAQ	0.652 ±0.350	0.510 ±0.237	< 0.001
Disease activity (according to DAS28)			
Remission	5 (3.6%)	9 (6.4%)	< 0.001
Low disease activity	8 (5.7%)	28 (20%)	
Moderate disease activity	39 (27.9%)	86 (61.4%)	
High disease activity	88 (62.9%)	17 (12.1%)	
Grades of synovitis according to ultrasound			
Right wrist	1.96 ±0.93	1.16 ±0.86	< 0.001
Left wrist	1.94 ±0.91	1.17 ±0.90	< 0.001
Right 2nd MCP	1.83 ±0.94	0.99 ±0.80	< 0.001
Left 2nd MCP	1.67 ±0.89	1.21 ±0.81	< 0.001
Right 3rd MCP	1.56 ±0.90	0.93 ±0.80	< 0.001
Left 3rd MCP	1.33 ±0.90	0.91 ±0.73	< 0.001
Right knee	0.89 ±0.81	0.54 ±0.68	< 0.001
Left knee	0.94 ±0.76	0.63 ±0.68	< 0.001
Total US score	13.56 ±5.18	8.02 ±4.28	< 0.001

CDAI – Clinical Disease Activity Index, DAS28 – Disease Activity Score based on 28 joints, ESR – erythrocyte sedimentation rate, mHAQ – multidimensional Health Assessment Questionnaire, MCP I – metacarpophalangeal.

moderate cases, where the affected joints may be totally or partially outside joints specified in the scale.

In the present study, the U9 ultrasonographic scale was evaluated to give the evaluator the ability to add the joint or tendon most clinically affected to the fixed set of eight joints.

Tendon pathology is an important feature in RA patients. Tendonitis may be a part of the inflammatory in RA and can lead to severe functional impairment in RA patients [18, 19].

One potential advantage of the U9 scale is that tendonitis is taken into account; also convergent and discriminant validities are two fundamental aspects of construct validity [20].

Convergent validity refers to the extent to which the new scale is related to other variables and other measures of the same construct [21].

The U9 scale in this study was shown to be significantly related to clinical parameters of disease activity such as DAS28-ESR and CDAI. Furthermore, U9 showed a high degree of correlation with the functional status as detected by HAQ.

Discriminant validity tests whether variables that are not supposed to be related are actually unrelated [22]. The current results demonstrated the ability of the U9 scale to distinguish different grades of RA activity and present well-defined cut-off points of different grades with very good specificity and sensitivity.

In the last decade, the use of musculoskeletal US has frequently been reported as a change-sensitive method for monitoring response to therapy in RA [23].

The present US study focused on the evaluation of synovial inflammatory alterations through treatment.

This study attempted to test the ability of the U9 scale to detect any change with treatment (i.e. decrease, increase or stable disease activity); however, this is not attributed to the efficacy of a specific type of therapy, as the treating physicians were free to prescribe any treatment to the patients studied.

The current study demonstrated that the change of the U9 US scale by treatment is correlated with changes in clinical and functional disease activity indices; thus the U9 scale can be useful for monitoring response to treatment in RA patients.

A prospective head-to-head study should be performed to compare performance of the U9 ultrasound scale with other composite scores such as U7 [24] and U8 [10] scales.

Limitations of the study

The main limitation of the present study is that a single sonographer in each center performed ultrasound assessments. It was not possible to assess inter-reader reliability.

Moreover, a head-to-head comparative study with the U7 score is warranted to test the sensitivity of this score in relation to other scores.

Conclusions

The U9 ultrasound scale has demonstrated correlations with clinical and functional scales. In addition, it can distinguish different grades of activity with well-defined cut-off values. The U9 scale can be used to monitor the therapeutic response in RA.

_____

References

1. Smolen JS, Aletaha D. The assessment of disease activity in rheumatoid arthritis. Clin Exp Rheumatol 2010; 28 (3 Suppl 59): S18–S27.
2. van der Heijde DM, van't Hof MA, van Riel PL, 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: 916–920, DOI: 10.1136/ard.49.11.916.
3. Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol 2005; 23 (5 Suppl 39): S100–S108.
4. Smolen JS, Steiner G. Therapeutic strategies for rheumatoid arthritis. Nat Rev Drug Discov 2003; 2: 473–488, DOI: 10.1038/nrd1109.
5. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995; 38: 727–735, DOI: 10.1002/art.1780380602.
6. Felson DT, Smolen JS, Wells G, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Ann Rheum Dis 2011; 70: 404–413, DOI: 10.1136/ard.2011.149765.
7. Slama IB, Allali F, Lakhdar T, et al. Reliability and validity of CDAI and SDAI indices in comparison to DAS28 index in Moroccan patients with rheumatoid arthritis. BMC Musculoskelet Disord 2015; 16: 268, DOI: 10.1186/s12891-015-0718-8.
8. Karimzadeh H, Rafiei R, Sayedbonakdar Z, Karami M. Agreement of clinical examination and ultrasound methods for detection of joints involvements in rheumatoid arthritis. J Res Med Sci 2017; 22: 87, DOI: 10.4103/jrms.JRMS_633_16.
9. Mortada, MA, Ghada G, Ebaid A, et al. FRI0624 U9: a novel clinically oriented ultrasonographic scale for assessing disease activity in rheumatoid arthritis. Ann Rheum Dis 2019; 78 (Suppl 2): 1009, DOI: 10.1136/annrheumdis-2019-eular.2505.
10. Yoshimi R, Ihata A, Kunishita Y, et al. A novel 8-joint ultrasound score is useful in daily practice for rheumatoid arthritis.

Mod Rheumatol 2015; 25: 379–385, DOI: 10.3109/14397595. 2014.974305.

11. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford) 2012; 51 (Suppl 6): vi5–vi9, DOI: 10.1093/rheumatology/kes279.

12. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. Health Qual Life Outcomes 2003; 1: 20, DOI: 10.1186/1477-7525-1-20.

13. Backhaus M, Burmester GR, Gerber T, et al. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001; 60: 641–649, DOI: 10.1136/ard.60.7.641.

14. 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, DOI: 10.1136/rmdopen-2016-000427.

15. Hammer HB, Bolton-King P, Bakkeheim V, et al. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. Ann Rheum Dis 2011; 70: 1995–1998, DOI: 10.1136/ard.2011.152926.

16. Mandl P, Naredo E, Wakefield RJ, et al. 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, DOI: 10.3899/jrheum.110424.

17. Mortada MA, Eitta H, Elmallah R, et al. AB1119 U9. A novel clinically oriented ultrasonographic scale as a useful marker for monitoring therapeutic response in rheumatic arthritis (multi

centers study). Ann Rheum Dis 2020; 79 (Suppl 1): 1848–1849, DOI: 10.1136/annrheumdis-eular2020.2733.

18. Janta I, Stanciu D, Hinojosa M, et al. Structural damage in rheumatoid arthritis: comparison between tendon damage evaluated by ultrasound and radiographic damage. Rheumatology (Oxford) 2016; 55: 1042–1046, DOI: 10.1093/rheumatology/kew020.

19. Elsaman AM, Mostafa ES, Radwan AR. Ankle evaluation in active rheumatoid arthritis by ultrasound: a cross-sectional study. Ultrasound Med Biol 2017; 43: 2806–2813, DOI: 10.1016/j.ultrasmedbio.2017.08.928.

20. Siemons L, Ten Klooster PM, Taal E, et al. Modern psychometrics applied in rheumatology – a systematic review. BMC Musculoskelet Disord 2012; 13: 216, DOI: 10.1186/1471-2474-13-216.

21. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959; 56: 81–105.

22. Farrell AM. Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). J Bus Res 2010; 63: 324–327, DOI: 10.1016/j.jbusres.2009.05.003.

23. do Prado AD, Staub HL, Bisi MC, et al. Ultrasound and its clinical use in rheumatoid arthritis: where do we stand? Adv Rheumatol 2018; 58: 19, DOI: 10.1186/s42358-018-0023-y.

24. Backhaus TM, Ohrndorf S, Kellner H, et al. The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. Ann Rheum Dis 2013; 72: 1163–1169, DOI: 10.1136/annrheumdis-2012-201397.