Arthritis Care & Research
Vol. 65, No. 6, June 2013, pp 879–887
DOI 10.1002/acr.21913

ORIGINAL ARTICLE

Clinical and Ultrasound-Based Composite Disease Activity Indices in Rheumatoid Arthritis: Results From a Multicenter, Randomized Study

P. MANDL,[1] P. V. BALINT,[2] Y. BRAULT,[3] M. BACKHAUS,[4] M. A. D'AGOSTINO,[5] W. GRASSI,[6] D. VAN DER HEIJDE,[7] E. DE MIGUEL,[8] R. J. WAKEFIELD,[9] I. LOGEART,[3] AND M. DOUGADOS[10]

Objective. **To evaluate the metrologic properties of composite disease activity indices in rheumatoid arthritis (RA), utilizing information derived from clinical, gray-scale (GS), and power Doppler (PD) ultrasound examinations, and to assess the classification of patients according to disease activity using such indices.**

Methods. **This ancillary study utilized data from a multicenter, prospective, randomized, parallel-group study conducted in subjects with moderate RA randomized to receive etanercept and methotrexate (ETN + MTX) or usual care (various disease-modifying antirheumatic drugs [DMARDs]). In multimodal indices, the 28 swollen joint count was either supplemented or replaced by clinically nonswollen joints in which the presence of synovitis was detected either by GS and/or PD and was calculated according to the Disease Activity Score in 28 joints (DAS28) or the Simplified Disease Activity Index (SDAI). Reliability, external validity, and discriminative capacity were calculated at baseline/screening by intraclass correlation coefficient, Pearson's correlation, and standardized response mean, respectively.**

Results. **Data from 62 patients (mean ± SD age 53.8 ± 13.2 years, mean ± SD disease duration 8.8 ± 7.7 years, mean ± SD disease activity 4.6 ± 0.5 [DAS28] and 20.9 ± 5.9 [SDAI]) were analyzed, with 32 receiving ETN + MTX and 30 receiving DMARDs. The metrologic properties were at least as good for GS- and/or PD-based indices as for their clinical counterparts. Using GS- and PD-supplemented indices, an additional 67.8% and 32.3% of patients (DAS28-derived and SDAI-derived indices, respectively) could be classified as having high disease activity at the screening visit.**

Conclusion. **Multimodal indices incorporating ultrasound and clinical data had similar metrologic properties to their clinical counterparts; certain indices allowed for a significantly larger number of patients to be classified to either high or moderate disease activity at the screening visit.**

INTRODUCTION

Composite or pooled indices are useful tools that combine core set variables for disease activity that have been recommended by international and national organizations to be used in rheumatoid arthritis (RA) disease activity assessment (1–3). Composite indices are used for the evaluation and quantification of disease activity in patients with RA both in everyday practice and in clinical trials (4–7) and contain a pooled set of different measures, including joint counts, patient and physician global assessments, and acute-phase reactants. All currently used composite disease activity indices include a swollen joint

ClinicalTrials.gov identifier: NCT00706797.

Supported by Wyeth Pharmaceuticals, which was acquired by Pfizer in October 2009.

[1]P. Mandl, MD, PhD: Medical University of Vienna, Vienna, Austria, and National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; [2]P. V. Balint, MD, PhD, FRCP: National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; [3]Y. Brault, MSc, I. Logeart, MD: Pfizer, Paris, France; [4]M. Backhaus, MD, PhD: University Hospital Charité, Berlin, Germany; [5]M. A. D'Agostino, MD, PhD: Versailles-Saint Quentin en Yvelines University, AP-HP, Ambroise-Paré Hospital, Boulogne-Billancourt, France; [6]W. Grassi, MD, PhD: Università Politecnica delle Marche, Jesi, Ancona, Italy; [7]D. van der Heijde, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [8]E. de Miguel, MD, PhD: La Paz University Hospital, Madrid, Spain; [9]R. J. Wakefield, MD, FRCP: University of Leeds and Chapel Allerton Hospital, Leeds, UK; [10]M. Dougados, MD: Paris-Descartes University, UPRES-EA 4058, AP-HP, Cochin Hospital, Paris, France.

Drs. Mandl and Balint contributed equally to this work.

Dr. Mandl has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from Pfizer, Wyeth, and Abbott. Dr. Balint has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from Pfizer, Abbott, MSD, Roche, UCB, Esaote, GE, Philips, Sonosite, the European League Against Rheumatism, and the American College of Rheumatology. Dr. Backhaus has

Significance & Innovations

- Multimodal disease activity indices are suitable instruments for measuring disease activity in patients with rheumatoid arthritis.
- Incorporating ultrasound data into composite disease activity indices may influence the classification of patients according to disease activity.

count (SJC) and a tender joint count, which are based on clinical examination.

The widespread use of musculoskeletal ultrasound by rheumatologists in both clinical practice and clinical trials has led to the development of various quantification methods, the majority of which were developed for the detection of synovitis in RA. As compared to clinical examination, ultrasound has demonstrated superior sensitivity and interobserver reliability with respect to the detection of synovitis in RA in a number of studies (8–12).

In order to quantify the pathologic changes in single joints, synovitis may be evaluated by binary grading (presence/absence), semiquantitative scoring (scales), and quantitative measurements (volume/depth of synovial tissue), both in gray-scale (GS) (8,13–15) and also by using color/power Doppler (PD) flow assessment (16–23). In recent years, several global ultrasound synovitis scoring systems that aim to assess synovitis on the patient level as opposed to the level of single joints have been developed and validated (24). Scores based on a reduced number of joints have shown a similar correlation with clinical and laboratory parameters as the corresponding 60-joint evaluation (25–27). A study evaluating a large number of such scoring systems has demonstrated that both binary and graded ultrasound-based global joint scoring systems have at least as good metrologic properties as their respective counterparts based on clinical examination using the Outcome Measures in Rheumatology (OMERACT) filter (28). We have recently published results from an ancillary ultrasound study to a multicenter, randomized controlled trial in which GS- and PD-based synovitis scoring systems demonstrated better reliability than generally used clinical indices for evaluating synovitis in RA, with PD also demonstrating at least as good discriminant capacity as clinical examination for distinguishing between treatment arms (29).

We have performed this ancillary study using data derived from the same multicenter, randomized controlled trial (29) to answer a different question regarding the incorporation of ultrasound-derived data into composite disease activity indices. To our knowledge, this is the first study to systematically evaluate multimodal disease activity indices that combine data derived from both clinical and ultrasound examination for measuring disease activity in RA using data from an interventional clinical trial. Recently, Damjanov et al published a study wherein they evaluated a composite disease activity index in which the clinical joint counts were replaced by GS and PD counts for 28 and 22 joints, respectively (30). Rather than replacing clinical data with ultrasound as seen in this latter study, our multimodal indices combine clinical and ultrasound data. The multimodal indices were based on available and widely utilized clinical indices for disease activity: the Disease Activity Score in 28 joints (DAS28) (5) and the Simplified Disease Activity Index (SDAI) (6). In the multimodal indices, the 28SJC was either supplemented or replaced by joints deemed nonswollen on clinical examination in which the presence of synovitis was detected either by GS and/or PD. In addition, the classification of patients according to disease activity was also assessed using the various indices.

MATERIALS AND METHODS

The etanercept (ETN) versus disease-modifying antirheumatic drugs (DMARDs) multicenter, prospective, randomized trial was a 52-week, multicenter, prospective, open-label, randomized, parallel-group, outpatient study in RA subjects with moderate disease activity (MDA). Inclusion and exclusion criteria as well as additional information on the trial have been reported previously (29). Briefly, in arm A, patients received once-weekly subcutaneous injections of 50 mg of ETN while continuing their current dose of methotrexate (MTX), while in the other, usual care arm, investigators prescribed DMARDs from a list of the 6 most commonly prescribed DMARDS in the participating countries. Subjects were scheduled to participate in the study for approximately 62 weeks, including a screening period of up to 6 weeks and an open-label treatment period of 52 weeks. Both clinical and ultrasound evaluations (see below) were performed on all participating patients at the screening and baseline visits, as well as at weeks 4, 12, 24, and 52. Safety was assessed throughout the course of the study. This study was approved by the appropriate ethical committees. All patients gave their written informed consent before entering the study.

Ultrasound was an integral but optional part of the trial from its initial conception. Centers participating in the

received a scientific grant from Pfizer. Dr. Grassi has received consultant fees and/or speaking fees (less than $10,000 each) from Abbott, Amgen, BMS, Esaote, GE, Menarini, MSD, Pfizer, Savient, Schering-Plough, Roche, UCB, and Wyeth. Dr. van der Heijde has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from Abbott, Amgen, AstraZeneca, BMS, Centocor, Chugai, Eli Lilly, GSK, Merck, Novartis, Otsuka, Pfizer, Roche, Sanofi-Aventis, Schering-Plough, UCB, and Wyeth. Dr. de Miguel has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from AbbVie, Pfizer, Menarini, BMS, MSD, Schering-Plough, Roche, UCB, Wyeth, and Abbott. Dr. Logeart owns stock and/or stock options in Pfizer. Dr. Dougados has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from Pfizer and Abbott.

Address correspondence to P. Mandl, MD, PhD, 20-21 Währinger Gürtel, 1080 Vienna, Austria. E-mail: mandl peter@yahoo.com.

Submitted for publication May 25, 2012; accepted in revised form November 6, 2012.

study had the option to participate in a preplanned optional ancillary ultrasonography study if they had access to any one of the high-end ultrasound machines (see below) as well as to an experienced sonographer, and possessed the facilities permitting an ultrasound evaluation totally independent from the clinical examination. In such centers, ultrasound evaluations were conducted on every patient included in the main study. The sponsor decided to terminate the original study, due to inadequate enrollment, before completion of the enrollment of the total sample size (700 subjects). Few subjects had completed the 52-week followup at the time of the study termination. Consequently, the analyses with respect to the ancillary study were restricted to data collected from screening to the week 12 visit period for which sufficient data were available to perform the evaluation.

Patients. MDA was defined by a DAS28 score of >3.2 and ≤ 5.1. Subjects with RA were enrolled in approximately 100 sites over 14 countries in Europe. The intent-to-treat (ITT) population for the ancillary study included all randomized subjects participating in the study.

Clinical evaluation. Clinical joint counts were performed by an investigator on the 28 joints included in the DAS28 (bilateral wrist, metacarpophalangeal joints 1–5, interphalangeal joint, proximal interphalangeal joints 2–5, elbow, shoulder, and knee). The investigator performing the clinical evaluation was unaware of the result of the ultrasonographic examination throughout the study. Joint swelling and tenderness were noted for each. The DAS28 and SDAI indices were calculated as shown in Table 1.

Ultrasound evaluation. Sonographers were blinded to the clinical status (including joint count status) of the patients throughout the study. Sonographers were not aware of the results of the clinical and laboratory examination as well as the drug allocation or regimen of the patients throughout the study. Ultrasound evaluation was performed on the same 28 joints assessed clinically. Systematic multiplanar GS and PD examination was carried out with commercially available real-time scanners (e.g., Esaote MyLab70 XVG, Esaote Technos MPX, General Electric Logiq 9, etc.) using multifrequency linear transducers (6–18 MHz). Ultrasonographic scanning techniques, GS and PD machine settings, and definitions of abnormality were standardized among the investigators prior to the study during a 2-day meeting. The ultrasound scanning method has been described previously (31). Sonographers were allowed to modify the machine settings (e.g., gain, pulse repetition frequency, etc.) on their individual machines to produce the best-quality images, allowing them to appropriately score each image. Synovitis was defined according to the published OMERACT definitions (32). Both GS and PD examinations were recorded for each of the 28 joints included in the DAS28. For each joint, the following regions were evaluated bilaterally: dorsal and volar aspects of the wrist (both longitudinal and transverse scans of the radiocarpal, ulnar-carpal, and radioulnar joints); dorsal and volar aspects of metacarpophalangeal

Table 1. Composition of composite disease activity indices*

	Composition
DAS28-derived indices	
DAS28 clinical	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}(28SJC) + 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 GS	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(GS-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 PD	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(PD-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 GS and PD	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(GS- and PD-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 + GS	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(28SJC + nonswollen 28 but GS-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 + PD	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(28SJC + nonswollen 28 but PD-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
DAS28 + GS/PD	$0.56 \times \sqrt{}(28TJC) + 0.28 \times \sqrt{}$(28SJC + nonswollen 28 GS- and/or PD-positive joints) $+ 0.70 \times \ln(ESR) + 0.014 \times GH$
SDAI-derived indices	
SDAI clinical	28TJC + 28SJC + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI GS	28TJC + 28 (GS-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI PD	28TJC + 28 (PD-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI GS and PD	28TJC + 28 (GS- and PD-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI + GS	28TJC + 28 (28SJC + nonswollen GS-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI + PD	28TJC + 28 (28SJC + nonswollen PD-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP
SDAI + GS/PD	28TJC + 28 (28SJC + nonswollen GS- and/or PD-positive joints) + PGA (0–10-cm VAS) + MDGA (0–10-cm VAS) + CRP

* DAS28 = Disease Activity Score in 28 joints; 28TJC = 28 tender joint count; 28SJC = 28 swollen joint count; ESR = erythrocyte sedimentation rate; GH = patient general health; GS = gray-scale; PD = power Doppler; SDAI = Simplified Disease Activity Index; PGA = patient global assessment; VAS = visual analog scale; MDGA = physician global assessment; CRP = C-reactive protein.

joints 1–5 (both longitudinal and transverse scans); dorsal and volar aspects of the interphalangeal joint (both longitudinal and transverse scans); dorsal and volar aspects of proximal interphalangeal joints 2–5 (both longitudinal and transverse scans); anterior aspect (longitudinal humeroradial, longitudinal humeroulnar, and anterior transverse scans) and posterior aspect of the elbow (both longitudinal and transverse scans); anterior aspect (biceps tendon sheath longitudinal and transverse scans), posterior aspect (transverse glenohumeral scan), and axillar aspect of the shoulder (longitudinal scan of the axillar recess); and anterior aspect of the knee (longitudinal and transverse scan of the suprapatellar recess). Joints demonstrating the presence of synovial thickening in GS were included in the GS synovitis joint count. Small joints in which the PD signal representing 1 or 2 vessels or more (including 1 confluent vessel) and large joints in which the PD signal representing 2 or 3 signals or more (including 2 confluent vessels) were demonstrated were included in the PD synovitis joint count.

Composite indices. A total of 14 different composite disease activity indices (Table 1) were evaluated, including the classic composite indices DAS28 and SDAI and multimodal indices based on clinical and GS and PD ultrasound–derived data, and calculated according to either the DAS28 or SDAI indices. In the first group of multimodal indices, the SJC was either replaced by the GS synovitis joint count or the PD synovitis joint count (DAS28 GS, DAS28 PD, SDAI GS, SDAI PD), or replaced by joints with a grade of ≥1 for both GS and PD (DAS28 GS and PD, SDAI GS and PD). In the second group, the SJC was supplemented also by those joints deemed nonswollen on clinical examination with a grade of ≥1 for GS and PD evaluation independently (DAS28 + GS, DAS28 + PD, SDAI + GS, SDAI + PD), or supplemented also by those joints deemed nonswollen on clinical examination with a grade of ≥1 for either GS or PD evaluation independently (DAS28 + GS/PD, SDAI + GS/PD).

Statistical analysis. Intraobserver reliability of each scoring system was calculated using screening and baseline visits on stable subjects. A difference of ≤10 on a 0–100 visual analog scale of the patient's and physician's global assessments of disease activity between the 2 consecutive visits was adopted as a definition for stable subjects. Intraobserver reliability was analyzed by the intraclass correlation coefficient (ICC) and its 2-sided 95% confidence interval (95% CI).

External validity, defined as the degree of association between the C-reactive protein (CRP) level and the various disease activity indices, was assessed at the screening visit using Pearson's correlation coefficient. The 2-sided 95% CI of this correlation coefficient is also provided using Fisher's z transformation.

The primary purpose of the study was to compare the discriminant capacity between the multimodal disease activity indices and their clinical counterparts, and not to check whether there was a statistically significant difference (e.g., a P value less than 0.05) between the treatment groups. Therefore, to calculate the discriminant capacity of the disease activity indices, the ETN + MTX and the conventional DMARD groups were evaluated separately using the baseline and week 12 visits, and the difference in standardized response mean (SRM) and its 2-sided 95% CI was calculated using the bootstrap resampling methodology. The SRM is calculated by dividing the mean change by the SD of the change score.

To classify the disease activity of patients using the DAS28-derived multimodal disease activity indices, the following cutoffs were employed using the original, SJC-based DAS28: low disease activity (LDA) = DAS28 ≤3.2, MDA = DAS28 >3.2 and ≤5.1, and high disease activity (HDA) = DAS28 >5.1. To classify the disease activity of patients using the SDAI-derived multimodal disease activity indices, the following cutoffs were employed using the original, SJC-based SDAI: LDA = SDAI ≤20, MDA = SDAI >20 and ≤40, and HDA = SDAI >40. All statistical analyses were performed using SAS, release 9.2.

RESULTS

In 18 centers located in 9 European countries, 66 patients were randomized in the ETN versus DMARDs multicenter, prospective trial before the trial was terminated by the sponsor due to inadequate enrollment. Sixty-two patients were included in the modified ITT population, with 30 patients receiving various DMARDs and 32 receiving ETN + MTX. One patient had a DAS28 score <3.2, while another had a score >5.1, at screening. Patient characteristics for the modified ITT population at screening are shown in Table 2.

Reliability. Intraobserver reliability of the composite synovitis scoring systems is shown in Table 3. Intraobserver reliability was assessed using screening and baseline values from 21 patients whose disease was con-

Table 2. Patient characteristics at screening*

	Value
Age, mean ± SD years	53.8 ± 13.2
Female sex, no. (%)	50 (80.6)
Disease duration, median (IQR) years	6.5 (11)
RF positivity (IU/ml), no. (%)	62 (100)
ACPA positivity (IU/ml), no. (%)	59 (95.1)
DAS28 (at baseline), mean ± SD	4.6 ± 0.5
SDAI (at baseline), mean ± SD	20.9 ± 5.9
HAQ DI score (at baseline), median (IQR)	1.35 (0.57)
ESR (at screening), median (IQR) mm/hour	28 (21)
CRP level (at screening), median (IQR) mg/liter	13.2 (12.4)

* IQR = interquartile range; RF = rheumatoid factor; ACPA = anti–citrullinated protein antibody; DAS28 = Disease Activity Score in 28 joints; SDAI = Simplified Disease Activity Index; HAQ = Health Assessment Questionnaire; DI = disability index; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein.

Table 3. Psychometric properties of composite disease activity indices*

	Intraobserver reliability, ICC (95% CI)	External validity, correlation with CRP level (95% CI)	Discriminant capacity, SRM (95% CI)
DAS28 clinical	0.52 (0.13, 0.77)	0.21 (−0.05, 0.44)	0.87 (0.28, 1.50)†
DAS28 GS‡	0.65 (0.31, 0.84)	0.24 (−0.02, 0.46)	0.85 (0.28, 1.45)†
DAS28 PD§	0.69 (0.38, 0.86)	0.23 (−0.02, 0.46)	0.80 (0.21, 1.43)†
DAS28 GS and PD¶	0.69 (0.38, 0.86)	0.24 (−0.02, 0.46)	0.79 (0.20, 1.43)†
DAS28 + GS#	0.74 (0.47, 0.89)	0.26 (0.00, 0.48)	0.70 (0.15, 1.44)†
DAS28 + PD**	0.72 (0.43, 0.88)	0.22 (−0.04, 0.45)	0.88 (0.33, 1.61)†
DAS28 + GS/PD††	0.75 (0.49, 0.89)	0.25 (−0.01, 0.47)	0.72 (0.15, 1.44)†
SDAI clinical	0.77 (0.51, 0.90)	0.41 (0.18, 0.61)	1.11 (0.25, 28.73)†
SDAI GS‡	0.87 (0.70, 0.94)	0.44 (0.20, 0.62)	1.09 (0.21, 51.16)†
SDAI PD§	0.89 (0.74, 0.95)	0.43 (0.20, 0.62)	1.17 (0.27, 74.24)†
SDAI GS and PD¶	0.89 (0.74, 0.95)	0.43 (0.20, 0.62)	1.17 (0.27, 74.24)†
SDAI + GS#	0.87 (0.70, 0.94)	0.42 (0.18, 0.61)	0.61 (−0.77, 3.43)
SDAI + PD**	0.86 (0.69, 0.94)	0.32 (0.06, 0.53)	0.52 (−0.51, 1.88)
SDAI + GS/PD††	0.90 (0.77, 0.96)	0.40 (0.16, 0.60)	0.84 (−0.24, 3.86)

* ICC = intraclass correlation coefficient; 95% CI = 95% confidence interval; CRP = C-reactive protein; SRM = standardized response mean; DAS28 = Disease Activity Score in 28 joints; GS = gray-scale; PD = power Doppler; SDAI = Simplified Disease Activity Index.
† Statistically significant intergroup difference.
‡ Clinical swollen joint count (SJC) replaced by swollen joints that also show signs of synovitis on GS.
§ Clinical SJC replaced by swollen joints that also show signs of synovitis on PD.
¶ Clinical SJC replaced by swollen joints that also show signs of synovitis either on GS and/or PD.
# Clinical SJC supplemented by nonswollen joints showing signs of synovitis on GS.
** Clinical SJC supplemented by nonswollen joints showing signs of synovitis on PD.
†† Clinical SJC supplemented by nonswollen joints showing signs of synovitis either on GS and/or PD.

sidered stable according to the definition described in the Methods. Reliability was found to be better for multimodal disease activity indices than for the composite indices based only on data from clinical evaluation of synovitis in stable subjects between the baseline and screening visits. Intraobserver reliability was ICC 0.71 (95% CI 0.65–0.75) for the multimodal DAS28-derived indices and ICC 0.52 for the clinical DAS28. Intraobserver reliability was ICC 0.88 (95% CI 0.86–0.90) for the multimodal SDAI-derived indices and ICC 0.77 for the clinical SDAI (Table 3).

External validity. External validity was evaluated based on the level of correlation existing between each index and the CRP level at screening. Overall correlation was weak among all scoring systems; however, GS- and PD-based composite scoring systems performed at least as well as their respective clinical counterparts.

External validity with respect to correlation with CRP level was 0.24 (95% CI 0.22–0.26) for the multimodal DAS28-derived indices and 0.21 for the clinical DAS28. Corresponding measures for the SDAI indices were 0.41 (95% CI 0.32–0.44) and 0.41, respectively (Table 3).

Discriminant capacity. Discriminant capacity was evaluated on the capacity of the changes in the different scoring systems to discriminate groups of patients (ETN versus usual care) between the baseline and week 12 visits. Discriminant capacity was SRM 0.79 (95% CI 0.70–0.88) for the multimodal DAS28-derived indices and SRM 0.87 for the clinical DAS28. Discriminant capacity was SRM 0.9 (95% CI 0.52–1.17) for the multimodal SDAI-derived indices and SRM 1.11 for the clinical SDAI (Table 3).

Although the objective of the substudy was not to compare discriminant capacity between the treatment groups (ETN versus usual care), it should be noted that the majority of indices, i.e., those with a 95% CI lowest limit above 0 (Table 3), were capable of demonstrating a statistically significant intergroup difference.

Classification of patients. There was no difference between the classic DAS28 and multimodal DAS28 indices in which the SJC was replaced by GS synovitis joint count or PD synovitis joint count (Figure 1) with regard to the classification of patients according to disease activity at screening. A decrease of 10% in the number of patients with MDA was revealed when using multimodal SDAI indices in which the SJC was replaced by PD synovitis joint count and GS synovitis joint count or PD synovitis joint count; these patients were reclassified as having LDA. However, indices in which the SJC was supplemented also by joints in which synovitis was detected by either GS, PD, or GS or PD ultrasound led to marked changes in classification, revealing increases in the number of patients classified as having HDA ranging from 45–67.8% and 13–32.3% (DAS28-derived and SDAI-derived indices, respectively) (Figure 1).

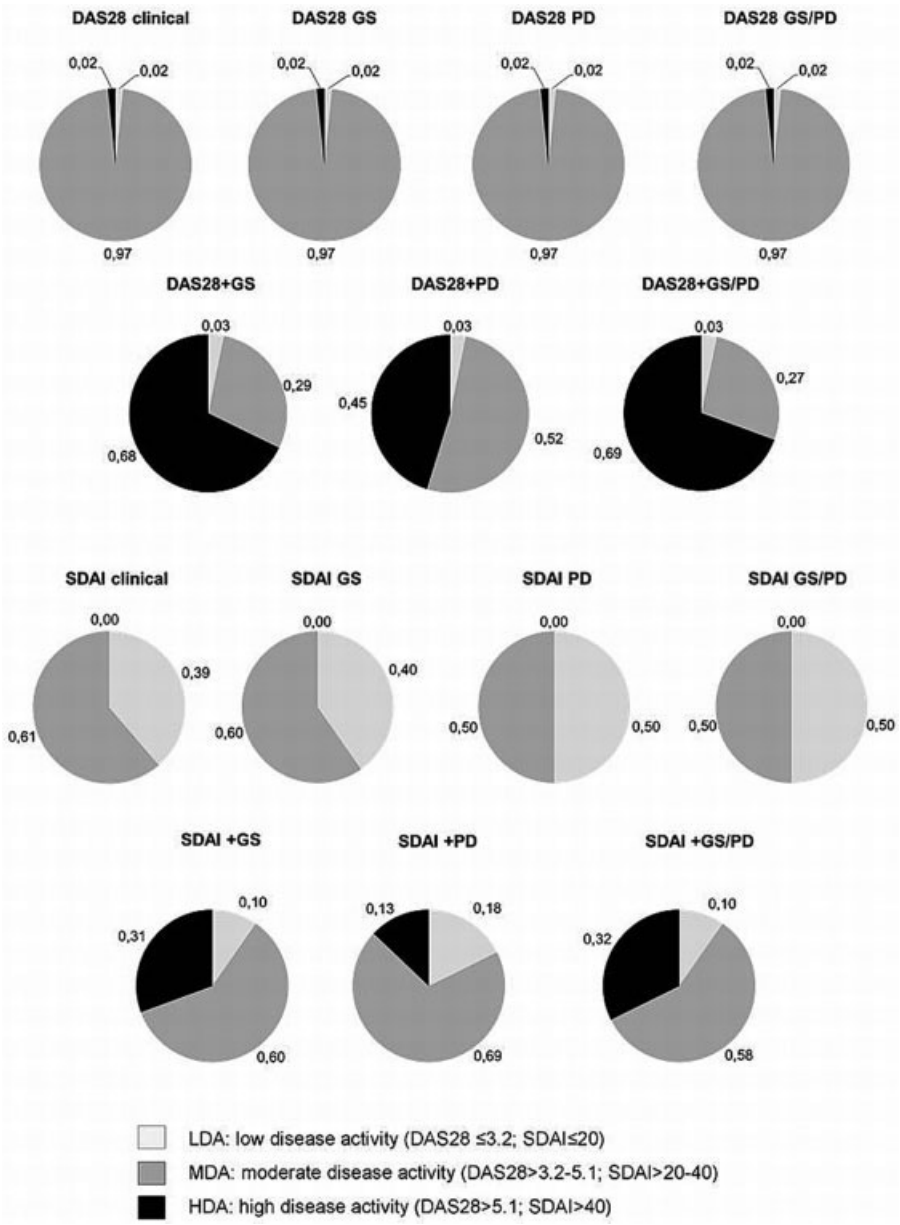

Figure 1. Classification of disease activity at screening using composite disease activity indices. DAS28 = Disease Activity Score in 28 joints; GS = gray-scale; PD = power Doppler; SDAI = Simplified Disease Activity Index.

DISCUSSION

The aim of this ancillary study was to evaluate the metrologic properties of composite disease activity indices using data derived from different evaluation modalities (clinical, GS, and PD) and assessed by a binary method for measuring synovitis in RA, and to assess the differences in classification when using these indices. Multimodal disease activity indices that combine data derived from both clinical and ultrasound examinations demonstrated better reliability as compared to their counterparts based solely on clinical examination.

Several studies have investigated the use of ultrasound in detecting and evaluating subclinical disease (33–37). Sonography detected synovitis as a surrogate of subclinical disease activity in almost two-thirds of patients with early, untreated oligoarthritis, whereas one-third of patients could be reclassified as having polyarticular disease (33).

The fact that no difference (DAS28 indices) or minimal difference (SDAI indices) could be observed in the distribution of patients according to disease activity when using multimodal indices in which the SJC was replaced by GS synovitis joint count or PD synovitis joint count (Figure 1)

might initially suggest that ultrasound confirms synovitis as assessed by GS or PD in the same joints that were deemed swollen on clinical examination. However, utilization of multimodal indices in which the SJC was supplemented also by joints in which synovitis was detected by either GS or PD ultrasound would have allowed an additional 67.8% and 32.3% of patients (DAS28-derived and SDAI-derived indices, respectively) to be classified as having HDA at the screening visit. Taken together, these 2 findings suggest a discrepancy between clinical and ultrasound examinations, in accordance with results from several studies (33–35). Supplementation of the SJC of the SDAI with ultrasound data increased disease activity; however, this increase was smaller in magnitude as compared to that seen with DAS28 indices. In addition, replacing the SJC with PD synovitis joint count or joints that were positive for both GS and PD in the SDAI led to a small decrease in disease activity, which was not seen with DAS28-derived indices. This is in line with studies showing evidence that the SDAI is a more stringent measure of remission because it allows for the least abnormalities of variables (38–40). These observations are also supported by a recent article by Balsa et al, who demonstrated that remission as classified by the SDAI is closer to the concept of an absence of inflammatory activity, when the absence of a positive PD signal is considered as a gold standard (41). In addition, the study by Saleem et al (42) showed that a larger percentage of patients in SDAI remission were in "imaging remission," as defined by the absence of GS synovitis and PD signal as compared to patients in DAS28 remission.

Whether the SDAI is "better" than the DAS28 cannot be fully answered based on the data from our study. Results from our study did, however, demonstrate higher overall metrologic properties for SDAI indices as compared to DAS28 indices. In particular, discriminative capacity was variable between the various indices, with SDAI indices in which the SJC was replaced by ultrasound-derived data outperforming those indices in which the SJC was supplemented. However, our study did not demonstrate any difference with respect to metrologic properties among the multimodal indices that would clearly designate any individual scoring system as being superior to its counterparts.

Our study has not revealed any advantages that would allow us to suggest the superiority of either GS or PD with regard to inclusion into composite indices. However, the reclassification data, specifically the DAS28 + GS and DAS28 + PD indices as well as the respective SDAI indices, show that a larger number of joints had GS signs of synovitis as compared to those that were positive for PD signal. When comparing classification according to DAS28 + GS/PD and SDAI + GS/PD with that of DAS28 + GS and DAS28 + PD, it becomes apparent that by supplementing a joint count that already includes both clinically swollen and GS-positive joints with joints that are positive for PD, there is only a very minor shift in disease activity.

The number of additional patients classified as having HDA with multimodal indices naturally decreased along with the progression of the study (to 11.3% and 3.2% by week 12) (data not shown).

Our study has several limitations. Early termination of the study meant that we could include a relatively small number of patients and were limited to perform analyses only of the baseline to 12-week period. In order to calculate intraobserver reliability, examinations had to be performed at 2 different time points, between which the disease activity, i.e., synovitis of the patients, could be reasonably expected to remain unchanged. We therefore chose the baseline and screening visits, during which no changes occurred in the patient's therapeutic regimen. We adopted a difference of $\leq$10 on a 0–100 visual analog scale of the patient's and physician's global assessments of disease activity as a proxy measure for stability. We are aware of the fact that although patient's and physician's global assessments of disease activity do not "guarantee" stable synovitis, they can nonetheless be considered a close indicator of stability based on the view of both the physician and the patient.

Patient's and physician's global assessments are components of the SDAI and the patient global assessment is a component of the DAS28; therefore, by limiting our assessment to patients in which these 2 measures remained relatively stable, we may have artificially reduced the variability of reliability. However, this would pose an overall effect and would still allow the comparison between multimodal and clinical indices. Despite the higher sensitivity of ultrasound for the detection of subclinical synovitis (33–35), multimodal indices showed better reliability than clinical indices. In our study, such indices failed to demonstrate superior discriminant capacity as compared to their clinical counterparts; however, since the responsiveness to change of a health status measurement instrument is closely related to its reliability (43), our data might suggest that multimodal indices, and indeed global synovitis scores based on ultrasound (28,29), which have been shown to be more reliable than their clinical counterparts, may potentially also be more responsive measures than clinical indices.

Feasibility data were not evaluated because we had information only on the time required to perform the 42-joint evaluation. In the present study, reliability was found to be higher for ultrasound, whereas discriminant capacities were similar for the techniques; this apparent discrepancy is explainable by the higher mean change observed in synovitis counts using ultrasound, as compared to the mean change observed clinically (data not shown). The capability of ultrasound to detect subclinical synovitis led to higher synovitis counts, but also to higher mean changes that, however, were similarly higher in both treatment groups (ETN versus usual care), which explains why the SRM values for the discriminant capacity of multimodal indices were similar to those for the clinical indices.

Our goal within this study was not to exchange clinical examination with ultrasound examination (30), but to combine the best of 2 diagnostic modalities, i.e., clinical examination and ultrasound. In this first systematic evaluation of multimodal disease activity indices combining both clinical and ultrasound data, such indices demonstrated better reliability as compared to their clinical counterparts and allowed for a significantly larger number

of patients to be reclassified to either HDA or MDA at the screening visit, therefore influencing eligibility for biologic therapy. What the actual impact and relevance of such changes in classification denote needs to be investigated and determined in further studies. A recent study comparing 9 conventional disease activity indices and the new American College of Rheumatology/European League Against Rheumatism remission criteria found no major differences among the evaluated indices in relation to physical functioning and radiographic progression as outcomes, despite marked differences in classification between the indices (40).

The improved sensitivity of novel imaging techniques has led to the notion of incorporating imaging into remission criteria for RA (42). As suggested by our study, incorporating imaging into measures of disease activity may improve the reliability of such systems and have marked effects on the classification of patients. This, however, required further evaluation and validation of multimodal disease activity indices in other studies.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Mandl had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Mandl, Balint, Brault, Backhaus, D'Agostino, Grassi, van der Heijde, de Miguel, Wakefield, Logeart, Dougados.

Acquisition of data. Mandl, Balint, Backhaus, Grassi, de Miguel, Wakefield, Dougados.

Analysis and interpretation of data. Mandl, Balint, Brault, Backhaus, D'Agostino, van der Heijde, Logeart, Dougados.

ROLE OF THE STUDY SPONSOR

Wyeth Pharmaceuticals, which was acquired by Pfizer Inc. in October 2009, was the sponsor of the trial. Two co-authors, Y. Brault, MSc, and I. Logeart, MD, are employees of Pfizer and participated as a statistician and clinical researcher, respectively, in the study design, analysis of the data, and the writing of the manuscript. The decision to submit the manuscript for publication and the publication of this article were not contingent upon approval by Wyeth Pharmaceuticals or Pfizer Inc.

REFERENCES

1. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729–40.
2. Smolen JS. The work of the EULAR Standing Committee on International Clinical Studies Including Therapeutic Trials (ESCISIT) [editorial]. Br J Rheumatol 1992;31:219–20.
3. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol Suppl 1994;41: 86–9.
4. 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.
5. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. 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.
6. 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 (Oxford) 2003;42:244–57.
7. 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:S100–8.
8. Backhaus M, Kamradt T, Sandrock D, Loreck D, Fritz J, Wolf KJ, et al. Arthritis of the finger joints: a comprehensive approach comparing conventional radiography, scintigraphy, ultrasound, and contrast-enhanced magnetic resonance imaging. Arthritis Rheum 1999;42:1232–45.
9. 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.
10. Szkudlarek M, Narvestad E, Klarlund M, Court-Payen M, Thomsen HS, Ostergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. Arthritis Rheum 2004;50:2103–12.
11. Karim Z, Wakefield RJ, Quinn M, Conaghan PG, Brown AK, Veale DJ, et al. Validation of ultrasonography in the detection of synovitis in the knee: a comparison with arthroscopy and clinical examination. Arthritis Rheum 2004;50:387–94.
12. Naredo E, Bonilla G, Gamero F, Uson J, Carmona L, Laffon A. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. Ann Rheum Dis 2005; 64:375–81.
13. 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.
14. 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.
15. Ribbens C, Andre B, Marcelis S, Kaye O, Mathy L, Bonnet V, et al. Rheumatoid hand joint synovitis: gray-scale and power Doppler US quantifications following anti-tumor necrosis factor-α treatment. Pilot study. Radiology 2003;229:562–9.
16. Szkudlarek M, Court-Payen M, Strandberg C, Klarlund M, Klausen T, Ostergaard M. 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.
17. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJ, Marsters PA, et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis. Arthritis Rheum 2004;50:1107–16.
18. Terslev L, Torp-Pedersen S, Savnik A, von der Recke P, Qvistgaard E, Danneskiold-Samsoe B, 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.
19. Qvistgaard E, Rogind H, Torp-Pedersen S, Terslev L, Danneskiold-Samsoe B, Bliddal H. Quantitative ultrasonography in rheumatoid arthritis: evaluation of inflammation by Doppler technique. Ann Rheum Dis 2001;60:690–3.
20. Teh J, Stevens K, Williamson L, Leung J, McNally EG. Power

Doppler ultrasound of rheumatoid synovitis: quantification of therapeutic response. Br J Radiol 2003;76:875–9.
21. Ozgocmen S, Kiris A, Kocakoc E, Ardicoglu O, Kamanli A. Evaluation of metacarpophalangeal joint synovitis in rheumatoid arthritis by power Doppler technique: relationship between synovial vascularization and periarticular bone mineral density. Joint Bone Spine 2004;71:384–8.
22. Kiris A, Ozgocmen S, Kocakoc E, Ardicoglu O. Power Doppler assessment of overall disease activity in patients with rheumatoid arthritis. J Clin Ultrasound 2006;34:5–11.
23. Carotti M, Salaffi F, Morbiducci J, Ciapetti A, Bartolucci L, Gasparini S, et al. Colour Doppler ultrasonography evaluation of vascularization in the wrist and finger joints in rheumatoid arthritis patients and healthy subjects. Eur J Radiol 2012;81:1834–8.
24. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA, on behalf of the OMERACT Ultrasound Task Force. A systematic literature review analysis of ultrasound joint count and scoring systems used to assess synovitis in rheumatoid arthritis, according to the OMERACT filter. J Rheumatol 2011;38:2055–62.
25. 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.
26. Naredo E, Rodriguez M, Campos C, Rodriguez-Heredia JM, Medina JA, Giner E, et al, for the Ultrasound Group of the Spanish Society of Rheumatology. 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.
27. Naredo E, Gamero F, Bonilla G, Uson J, Carmona L, Laffon A. Ultrasonography assessment of inflammatory activity in rheumatoid arthritis: comparison of extended versus reduced joint evaluation. Clin Exp Rheumatol 2005;23:881–7.
28. Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J, et al. Evaluation of several ultrasonography scoring systems of synovitis and comparison to clinical examination: results from a prospective multi-center study of rheumatoid arthritis. Ann Rheum Dis 2010;69:828–33.
29. 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.
30. Damjanov N, Radunovic G, Prodanovic S, Vukovic V, Milic V, Simic Pasalic K, et al. Construct validity and reliability of ultrasound disease activity score in assessing joint inflammation in RA: comparison with DAS-28. Rheumatology (Oxford) 2012;51:120–8.
31. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al, for the Working Group for Musculoskeletal Ultrasound in the EULAR Standing Committee on International Clinical Studies including Therapeutic Trials. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;60:641–9.
32. 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.
33. Wakefield RJ, Green MJ, Marzo-Ortega H, Conaghan PG, Gibbon WW, McGonagle D, et al. Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. Ann Rheum Dis 2004;63:382–5.
34. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, 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.
35. Magni-Manzoni S, Epis O, Ravelli A, Klersy C, Visconti C, Lanni S, et al. Comparison of clinical versus ultrasound-determined synovitis in juvenile idiopathic arthritis. Arthritis Rheum 2009;61:1497–504.
36. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Karim Z, et al. Disease remission state in patients treated with the combination of tumor necrosis factor blockade and methotrexate or with disease-modifying antirheumatic drugs: a clinical and imaging comparative study. Arthritis Rheum 2009;60:1915–22.
37. 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.
38. Aletaha D, Smolen JS. Remission of rheumatoid arthritis: should we care about definitions? Clin Exp Rheumatol 2006;24 Suppl:S45–51.
39. Mierau M, Schoels M, Gonda G, Fuchs J, Aletaha D, Smolen JS. Assessing remission in clinical practice. Rheumatology (Oxford) 2007;46:975–9.
40. Klarenbeek NB, Koevoets R, van der Heijde DM, Gerards AH, Ten Wolde S, Kerstens PJ, et al. Association with joint damage and physical functioning of nine composite indices and the 2011 ACR/EULAR remission criteria in rheumatoid arthritis. Ann Rheum Dis 2011;70:1815–21.
41. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. Rheumatology (Oxford) 2010;49:683–90.
42. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Wakefield R, et al. Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. Ann Rheum Dis 2011;70:792–8.
43. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res 2001;10:571–8.