Clin Rheumatol (2016) 35:467–471
DOI 10.1007/s10067-014-2849-y

BRIEF REPORT

Patient self-assessment of flare in rheumatoid arthritis: criterion and concurrent validity of the Flare instrument

Annette de Thurah · Thomas Maribo · Kristian Stengaard-Pedersen

Received: 26 October 2014 / Accepted: 7 December 2014 / Published online: 18 December 2014

Abstract The French Flare instrument (FI) aims to identify flares in rheumatoid arthritis between consultation. The objective of the present study was to present both concurrent and criterion validity of the Danish version of FI, as compared to DAS28-CRP. The study was a cross-sectional study comparing FI with DAS28-CRP among patients with rheumatoid arthritis (RA) in connection with the same outpatient visit. The study population consisted of 117 prevalent patients diagnosed with RA according to the ACR 1987/2010 criteria. Consecutive patients were included in the study in relation to their outpatient treatment at the Department of Rheumatology, Aarhus University Hospital, Denmark between 01 October 2012 and 31 December 2012. The sensitivity and specificity were 85.4 (95 % CI, 72.2; 93.9) and 50.7 (95 % CI, 38.4; 63.0), respectively. The positive predictive value was 53.6 (95 %CI, 47.0; 60.1) and the negative predictive value 83.9 (95 % CI, 71.7; 91.5). Positive and negative likelihood ratio were 1.73(95 % CI, 1.33; 2.26) and 0.29 (95 % CI, 0.14; 0.59). Tests with high sensitivity and small LR are most useful for ruling out the disease. Hence, our findings indicate that FI works well in ruling out a flare among patients with RA.

Keywords Flare · Flare instrument · Patient reported outcome measure · Rheumatoid arthritis · Validation

A. de Thurah (✉) · K. Stengaard-Pedersen
Department of Rheumatology, Aarhus University Hospital, Nørrebrogade 44, Aarhus C, 8000 Aarhus, Denmark
e-mail: annethur@rm.dk

T. Maribo
Rehabilitation Center Marselisborg, Department of Public Health, Section of Social Medicine and Rehabilitation, Aarhus University, Aarhus, Denmark

T. Maribo
CFK, Central Denmark Region, Aarhus, Denmark

Introduction

The newly developed Flare instrument (FI) is a patient self-assessment questionnaire aiming to detect changes in disease activity among patients with rheumatoid arthritis (RA) in daily clinical practice [1]. FI seeks to capture the patient perspective of a flare in RA as recommended by the Outcome Measures in Rheumatology (OMERACT) group. It comes within the scope of the term *patient-related outcome measure* (PROM), i.e., self-completed questionnaires that measure health status or health-related quality of life [2]. The ability of the instrument to actually capture a flare in RA depends, however, on the psychometric strength of the instrument. In RA, the gold standard for measuring disease activity is Disease Activity Score 28, CRP (DAS28-CRP) [3].

In a previous study, we have described the translation process into Danish and the reliability of the FI. Thus, the aim of the present study is to present both the concurrent and criterion validity of the FI as compared to DAS28-CRP.

Methods

The study is a cross-sectional study comparing FI with DAS28-CRP among patients with RA in connection with the same outpatient visit.

Patients

The study population consisted of 117 prevalent patients diagnosed with RA according to the ACR 1987/2010 criteria. Consecutive patients were included in the study in relation to their outpatient treatment at the Department of Rheumatology, Aarhus University Hospital, Denmark between 01 October 2012 and 31 December 2012.

The Flare instrument

The FI is a newly developed self-administrated tool designed to identify flares in RA between consultations by detecting both past and present disease activities in daily clinical practice. FI is a result of 99 semi-structured interviews together with statements from 13 rheumatologists generated through a Delphi process. It consists of 12 items, and each item represents a statement, which is related to disease activity in RA. FI addresses both the patients and the clinician's experience of disease activity [1].

When scoring FI, patients are asked to enumerate their degree of agreement on a 10-point Likert scale (0=completely agree, 10=completely disagree). The total score, Flare Total (FI Tot), is the arithmetic mean of the 12 items with potentially two subscales: one related to joint symptoms, Flare Joint (Fl Joint), and one related to general symptoms, Flare General (Fl Gen). Recently, FI has shown high reproducibility in a randomized controlled trail among 200 French RA patients with stable disease (unpublished data).

Criterion validity

Criterion validity concerns the degree to which the scores of an instrument are an adequate reflection of a "criterion standard" [4]. Thus, the criterion standard represents the best available instrument. In this analysis, we used DAS28-CRP as criterion standard, and furthermore, we compared the scores of the FI to Health Assessment Questionnaire (HAQ) and C-reactive protein (CRP) (concurrent validity).

Disease Activity Score 28

DAS28 is the most commonly used instrument to measure disease activity in RA and is recommended as the standard disease activity measure by the European League Against Rheumatism (EULAR) [3, 5, 6]. DAS28-CRP is a composite score including the patient's global assessment, report of physical functioning (HAQ), and the measurement of an acute phase reactant (CRP), together with a physician-based count of tender and swollen joints. The DAS28-CRP score runs from 0 to 9.4, and RA disease activity is defined as followed: DAS28-CRP$\leq$3.2: mild disease activity, DAS28-CRP >3.2 to $\leq$5.1: moderate disease activity, DAS28-CRP >5.1: high disease activity. Remission is defined as DAS28-CRP$\leq$2.6 [3].

Concurrent validity

The HAQ is the most widely used measure of functional disability in RA [7]. It is a validated measure, assessing difficulty over the past week in 20 specific functions grouped into eight categories: dressing and grooming, arising, eating, walking, personal hygiene, reaching, gripping, and other activities. The patients respond on a scale from 0 (no disability) to 3 (completely disabled).

C-reactive protein

CRP is a component in both the ACR criteria and the DAS28-CRP and is routinely used in clinical practice. CRP is considered the most useful biochemical marker for evaluating disease activity in patients with RA [8]. CRP measures are sensitive to short-term changes in disease activity and are one of the best indicators of the acute-phase response to inflammation [9].

Statistical analysis

Descriptive statistics were calculated for all variables. Criterion validity was evaluated using receiver operating characteristic (ROC) curves with DAS28-CRP as the external criterion, and the area under the curve was calculated. The optimal cutoff score for the FI compared to DAS28-CRP and the accompanying sensitivity, specificity rates, positive and negative predictive value, and likelihood ratio were determined. Cutoffs for the correlations y was evaluated by the criteria defined by Landis and Koch, who characterized values <0 as indicating no agreement and 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement [10].

Spearman's correlations were used as CRP, and HAQ was not normally distributed. Statistical analysis was performed with STATA13 software (StataCorp, College station).

Ethics

All patients gave their written informed consent. The study protocol was approved by the Danish Data Protection Agency (reference number: 1-16-02-577-13), and all procedures were in accordance with the declaration of Helsinki II.

Results

A total of 117 prevalent RA patients were included in the study. Demographic characteristics are shown in Table 1. All patients were treated with conventional DMARDs, and none of the patients received biological drugs.

Criterion validity

In the criterion validity test, we chose a cutoff of 2.5 on the FI Tot based on the receiver operating curve (ROC) (Fig. 1) and a cutoff of $\geq$3.2 on DAS28-CRP [3].

Table 1 Characteristics of 117 RA patients in the Flare validity study

Age in years	62.3 (12.1) [26–87]
Sex, m/f (%)	25/92 (21/79)
Rheuma factor-positive (%)[a]	81 (79.4)
Receiving corticosteroids (%)[b]	17 (15.3 %)
Receiving treatment with biological drugs (%)	0
DMARD (%)[c]	89 (76 %)

Values are mean (SD) and [range], *n* and (percentage) or median, and [interquartile range]

[a] $n=102$, due to missing data

[b] $n=111$, due to missing data

[c] Given as mono or combination therapy

In our cohort, the prevalence of DAS28-CRP ≥3.2 was approximately 40 %, resulting in a sensitivity (sens) and specificity (spec) of 85.4 % (95 % CI, 72.2; 93.9) and 50.7 % (95 % CI, 38.4; 63.0) and positive predictive (PP+) and negative predictive (PP−) values of 53.6 % (95 % CI, 47.0; 60.1) and 83.9 % (95 % CI, 71.7; 91.5) (Table 2). The area under the curve was approximately 77 %.

When adding CRP≥10 mmol/L to FI Tot≥2.5, the prevalence decreased to 20 % and the following diagnostic values were calculated: sens: 87.5 % (95 % CI, 67.6; 97.3) spec: 41.9 % (95 % CI, 31.8; 52.6), PP+: 27.4 % (95 % CI, 23.0; 32.2), PP−: 93.1 (95 % CI: 81.9; 97.5). Both models produced a negative likelihood ratio of approximately 0.30, providing moderate to strong evidence to rule out a flare (Table 2).

Concurrent validity

The Spearman correlation coefficients comparing FI Tot, Joint, and Gen with DAS28-CRP, CRP, and the HAQ are seen in Table 3. We detected a moderate correlation between DAS28-CRP and FI Tot, a poor correlation between CRP and FI Tot, and a poor correlation between FI Tot and the HAQ.

Discussion

The current study presents the diagnostic proprieties of the FI using DAS28-CRP as the criterion standard. The area under the curve is 77.2 %, indicating good to moderate diagnostic proprieties. Tests with high sensitivity and small LR are most useful for ruling out the disease, that is, a negative result indicates that the disease is not likely to be present [11]. Hence, our findings, showing a sensitivity of 85.4 % and a negative LR of 0.29, indicate that FI works well in ruling out a flare among patients with RA.

The correlation with DAS28-CRP is best for FI total compared with FI joint and FI general. The correlation with both CRP and HAQ is poor. We found a normal distribution of the DAS 28; however, the difference between calculating a Pearson instead of a Spearman correlation coefficient was insignificant.

Recently, the OMERACT group has published data on the validity of the OMERACT arthritis flare domains showing that PRO data can discriminate between patients with and without a worsening in RA disease activity, especially data on pain [12]. The present study is, however, the first study to investigate the validity of a flare instrument that integrates several PRO items into one score [1].

The current study has several potential limitations that merit further discussion. First, the construct of the FI Tot and

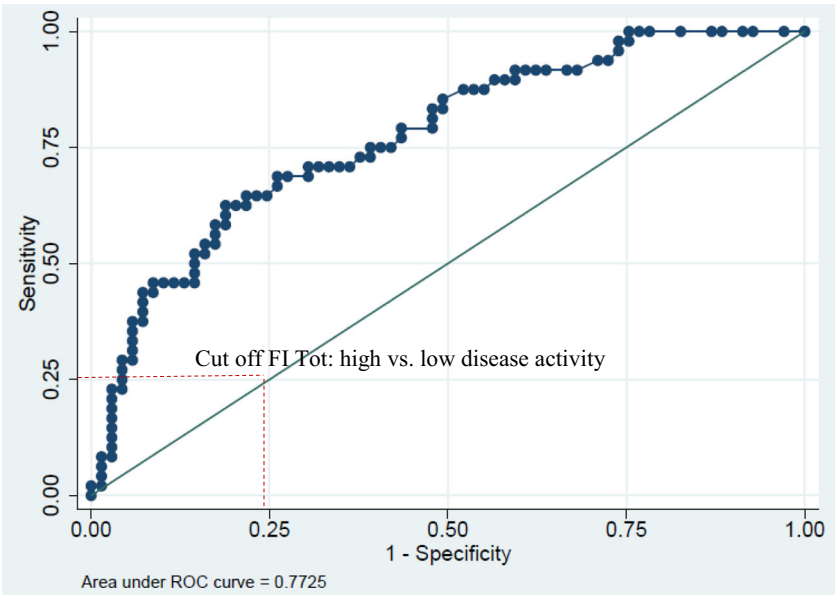

Fig. 1 Receiver operating characteristics (ROC) curve for sensitivity and 1-specificity for Fl Total≥2.5 and DAS28-CRP≥3.2 among 117 patients with rheumatoid arthritis, Aarhus University Hospital, Denmark Cutoff FI Tot: high vs. low disease activity

Table 2 Diagnostic test values of Flare Total (FI Tot) and FI Tot and elevated C-reactive protein using DAS28 as criterion standard, among 117 RA patients treated with conventional DMARDs, Aarhus University Hospital, Denmark

	Sensitivity (95 % CI)	Specificity (95 % CI)	Predictive value of positive test (95 % CI)	Predictive value of negative test (95 % CI)	Likelihood ratio + (95 % CI)	Likelihood ratio ÷ (95 % CI)
Fl Tot	85.4 % (72.2; 93.9)	50.7 % (38.4; 63.0)	53.6 %[a] (47.0, 60.1)	83.9 %[a] (71.7; 91.5)	1.73 (1.33; 2.26)	0.29 (0.14; 0.59)
Fl Tot and increased CRP[b]	87.5 % (67.6; 97.3)	41.9 % (31.8; 52.6)	50.1 %[c] (44.4; 55.8)	83.4 %[c] (63.0; 93.7)	1.51 (1.20; 1.90)	0.30 (0.10; 0.88)

[a] Prevalence of DAS28-CRP≥3.2: 40 %

[b] CRP≥10 mmol/L

[c] Prevalence of DAS28-CRP≥3.2 and CRP≥10 mmol/L: 20 %

the DAS28-CRP is not exactly the same. FI is intended to measure a change in disease activity between consultations [1] whereas the DAS28-CRP measures present disease activity [6]. Still, in daily clinical practice, DAS28-CRP is well established as a tool for monitoring the presence of a flare between consultations in the treat-to-target strategy [13]. Thus, it seems relevant to use DAS28-CRP as the criterion standard in the present study.

Second, the FI is measured on a continuous scale and has no a well-defined cutoff for high vs. low disease activity [1]. A full picture of the relationship between sensitivity and specificity can only be given from the ROC curve, and as can be seen, the FI was stable between 0 and 2.5, after which it increased. Based on this, we a priori choose a cutoff of 2.5.

The HAQ measures both past and present disease activities [7] and is known to measure functional disability rather than flare. Given this, HAQ measures a completely different construct than the FI, and therefore, our finding of a non-correlation between FI and HAQ supports the diagnostic abilities of the FI Tot in actually measuring a flare.

CRP is considered a useful biochemical marker for evaluating disease activity in RA [8]. However, an Austrian cohort study including 767 RA patients investigated the contribution of CRP to DAS28-CRP among patients with RA and found that 95 % of the variance of DAS28-CRP was explained without the inclusion of CRP [14]. This is in accordance with our findings that CRP added little to the FI Tot and that the correlation between FI tot and CRP was low. When using the FI in daily clinical practice, it is, however, important to keep in mind that the immunosuppressive medical treatment of RA influences the CRP level [15].

In conclusion, FI integrates the patients' perspective on flares as recommended by the OMERACT group [16] and further works well as a tool for ruling out flares in RA. Consequently, the FI has the potential of being a tool for decision aid, i.e., in a telemedicine intervention, and thus invites to new ways of organizing the continuous monitoring and treat-to-target strategy for patients with RA.

Table 3 Spearman correlation coefficients between Flare Total (FI Tot), Flare Joint (FI Joint) and Flare General (FI Gen) and DAS28-CRP, C-reactive protein, and the Health Assessment Questionnaire (HAQ) among 117 RA patients treated with conventional DMARDs, Aarhus University Hospital, Denmark

	DAS 28	*P* value	CRP	*P* value	HAQ	*P* value
Flare total	0.64	<0.001	0.27	0.01	0.48	<0.001
Flare joint	0.61	<0.001	0.29	0.005	0.39	<0.001
Flare general	0.56	<0.001	0.22	0.04	0.50	< 0.001

Acknowledgments We highly appreciate the collaboration with Professor Bruno Fautrel, Department of Rheumatology, Pitié-Salpétrière Hospital, Paris Cedex, France, who allowed us to translate the FI questionnaire from French into Danish and to carry out the first reliability and validity study. Financial support was provided by the Clinical Institute of Medicine, Aarhus University

Disclosures None

References

1. Berthelot JM, De Bandt M, Morel J et al (2012) A tool to identify recent or present rheumatoid arthritis flare from both patient and physician perspectives: the 'FLARE' instrument. Ann Rheum Dis 71:1110–1116
2. Department of Health (2010) Equity and excellence, Liberating the NHS. 2010 PROM data. Available at http://www.england.nhs.uk/statistics/statistical-work-areas/proms/
3. Dougados M, Aletaha D, van Riel P (2007) Disease activity measures for rheumatoid arthritis. Clin Exp Rheumatol 25:S22–9
4. Mokkink LB, Terwee CB, Patrick DL et al (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 19:539–549
5. Felson DT, Anderson JJ, Boers M et al (1993) The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 36:729–740

6. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL (1995) 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 38:44–48
7. Bruce B, Fries JF (2003) The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. J Rheumatol 30:167–178
8. Yildirim K, Karatay S, Melikoglu MA, Gureser G, Ugur M, Senel K (2004) Associations between acute phase reactant levels and disease activity score (DAS28) in patients with rheumatoid arthritis. Ann Clin Lab Sci 34:423–426
9. Kushner I (1991) C-reactive protein in rheumatology. Arthritis Rheum 34:1065–1068
10. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
11. Deeks JJ, Altman DG (2004) Diagnostic tests 4: likelihood ratios. BMJ 329:168–169
12. Lie E, Woodworth TG, Christensen R et al (2014) Validation of OMERACT preliminary rheumatoid arthritis flare domains in the NOR-DMARD study. Ann Rheum Dis 73:1781–1787
13. Smolen JS, Aletaha D, Bijlsma JW et al (2010) Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 69:631–637
14. Aletaha D, Nell VP, Stamm T et al (2005) Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 7:R796–806
15. Fleischmann R, van der Heijde D, Koenig AS et al (2014) How much does Disease Activity Score in 28 joints ESR and CRP calculations underestimate disease activity compared with the Simplified Disease Activity Index? Ann Rheum Dis 0:1–6
16. Kirwan JR, Ahlmen M, de Wit M et al (2005) Progress since OMERACT 6 on including patient perspective in rheumatoid arthritis outcome assessment. J Rheumatol 32:2246–2249