Changes in physical function using three methods of scoring the Health Assessment Questionnaire in established active rheumatoid arthritis

Winchow, L.L.; Tikly, M.; Musenge, E.; Chopra, A.; Huizinga, T.W.J.; Salomon-Escoto, K.; ... ; Govind, N.

Original Article

Changes in Physical Function using Three Methods of Scoring the Health Assessment Questionnaire in Established Active Rheumatoid Arthritis

Lai Ling Winchow, Mohammed Tikly, Eustasius Musenge[1], Arvind Chopra[2], T.W.J. Huizinga[3], Karen Salomon-Escoto[4], José Tavares-Costa[5], Nimmisha Govind

Division of Rheumatology, Department of Medicine, Chris Hani Baragwanath Academic Hospital, University of the Witwatersrand, [1]Division of Biostatistics and Epidemiology, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa, [2]Center for Rheumatic Diseases, Pune, Maharashtra, India, [3]Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands, [4]Division of Rheumatology, Department of Medicine, UMass Memorial Health and University of Massachusetts Medical School, Worcester, Massachusetts, United States of America, [5]Department of Rheumatology, Unidade Local de Saúde do Alto Minho, Ponte de Lima, Portugal

Received: 15-Apr-2022
Revised: 24-Jun-2022
Accepted: 05-Jul-2022
Published: 27-Aug-2022

Address for correspondence:
Dr. Lai Ling Winchow, Division of Rheumatology, Chris Hani Baragwanath Academic Hospital, P. O. 2013 Bertsham, South Africa.
E-mail: lai_ling.winchow@yahoo.com

Abstract

Background: We investigated sensitivity to change of three scoring methods of the Health Assessment Questionnaire (HAQ) in relation to change in disease activity in patients with active rheumatoid arthritis (RA).

Patients and Methods: Adult RA-patients with complete data in the Measurement of Efficacy of Treatment in the Era of Outcome in Rheumatology database with respect to the 20 HAQ questions and disease activity score with 28-joint count using the erythrocyte sedimentation rate (DAS28-ESR) for 2 visits, at least 6–12 months apart, and high disease activity (DAS28-ESR >5.1) at visit 1. Changes in HAQ scored by the (1) conventional method (HAQ-8), (2) HAQ-Tomlin method (HAQ-T), and (3) HAQ-20-item method (HAQ-20) were analyzed in relation to the European League Against Rheumatism (EULAR) RA response criteria, dichotomized to good/moderate and no response.

Results: In 421 patients, mean standard deviation (SD) DAS28-ESR declined significantly (6.1 [0.8]–4.8 [1.6], $P < 0.0001$), over a mean period (SD) of 8.7 (1.9) months. Median HAQ scores improved by all three scoring methods, HAQ-8 (1.6–1.4); HAQ-T (1.2–0.7); and HAQ-20 (1.2–0.9) with similar effect sizes of 0.97, 0.96, and 0.95, respectively. The proportion who achieved a HAQ minimally clinically important improvement (MCII) of ≥0.22 was significantly higher in 47% of patients with EULAR good/moderate score compared to the no response patients (64% vs. 11%, $P < 0.0001$). Good/moderate EULAR response, higher baseline DAS28, and higher baseline HAQ (7.11, 1.55, and 1.06, respectively) were independent predictors of achieving a HAQ-MCII.

Conclusion: Three HAQ scoring methods performed similarly in sensitivity to change with no advantage of alternative scoring methods compared to the conventional HAQ-8 method. A good/moderate EULAR response, despite long disease duration, was associated with a significant likelihood of achieving a HAQ-MCII.

Key Words: *Disease activity, health assessment questionnaire, Measurement of Efficacy of Treatment in the Era of Outcome in Rheumatology, rheumatoid arthritis*

Introduction

Rheumatoid arthritis (RA) is a chronic progressive inflammatory joint disorder that can result in irreversible joint damage and physical disability.[1] The modified Stanford Health Assessment Questionnaire (HAQ) is a validated and widely used disease-specific instrument for measuring functional disability in RA.[2] HAQ scores are influenced by a combination of joint inflammation (reversible) and damage (irreversible)[3] and are generally more sensitive to change in early RA than in established, long-standing RA.[4]

The standard method of scoring the HAQ (HAQ-8) is the sum of the worst scores in each of the 8 domains,

Access this article online	
Website: www.indianjrheumatol.com	**Quick Response Code**
DOI: 10.4103/injr.injr_80_22	

How to cite this article: Winchow LL, Tikly M, Musenge E, Chopra A, Huizinga TW, Salomon-Escoto K, *et al.* Changes in physical function using three methods of scoring the health assessment questionnaire in established active rheumatoid arthritis. Indian J Rheumatol 2023;18:S76-80.

consisting of at least two items, divided by 8.[5] A major drawback of this scoring method is the floor and ceiling effects. The only improvement in the worse item score results in an improvement in the domain score, and conversely, improvement in the item initially scoring worst may be masked by worsening of other items in the same domain. Hence, the instrument is less sensitive to change at the extremes of the scale, i.e., a patient with a HAQ of 0 seemingly cannot get better and, conversely, a patient with a HAQ of 3 cannot get worse.

Different methods of scoring the HAQ have been tried in an attempt to improve sensitivity to change. Tomlin *et al*. used the mean score per category to calculate the HAQ-Tomlin method (HAQ-T).[6] They showed a smaller variance in category scores, a better correlation between category scores and final score, and better reproducibility on repeated administrations of the questionnaire, compared to the HAQ-8 method. A rescored 20-item HAQ investigated by Wolfe showed better performance than the conventional HAQ in patient and item separation.[7]

The primary objective of the present study was to compare alternative HAQ scoring methods with the HAQ-8 method with respect to sensitivity to change in relation to disease activity in patients enrolled in a large multinational, the Measurement of Efficacy of Treatment in the Era of Outcome in Rheumatology (METEOR) database.[8] The secondary objective was to determine the proportion of patients achieving minimal clinical improvement of the HAQ scores by the three methods in relation to the European League Against Rheumatism (EULAR) RA clinical response criteria.

Patients and Methods

Adult patients fulfilling the 2010 ACR/EULAR RA classification criteria[9] in the METEOR database were studied. Only patients who had completed all 20 HAQ items and 4 variable disease activity score with 28-joint count using the erythrocyte sedimentation rate (DAS28-ESR)[10] data for any 2 visits at least 6–12 months apart, with high disease activity (DAS28-ESR >5.1) at the first of the two visits (henceforth termed visit 1) were included in the analysis [Supplementary Figure 1]. The rationale for including only patients with high disease activity was to enrich the sample to show changes in both disease activity and HAQ scores between two visits within a relatively short interval between visits. HAQ scores were analyzed in relation to EULAR RA response criteria, based on changes in DAS28-ESR scores.[11] Briefly, the HAQ was scored (range of 0–3) as follows: HAQ-8 as a quotient of the sum of worst scores per domain;[5] HAQ-T as a quotient of the sum of mean scores per domain divided by 8;[6] and HAQ-20-item method (HAQ-20) as a quotient of the sum of 20 items scores divided by 20. For the purpose of the study, patients were dichotomized into either EULAR good/moderate response or no response. The study was approved by the University of the Witwatersrand Human Research Ethics Committee (clearance certificate number M140950) on 20/02/2015.

Statistical method

A Bland–Altman plot was created to assess the agreement between scoring methods.[12,13] Effect sizes were calculated using the Guyatt's and Cohen's d methods.[14,15] The minimally clinically important difference ≥0.22[16] in HAQ score was used to define minimally clinically important improvement (MCII). Cumulative probability curves of the change in HAQ scores by the three methods were plotted for the good/moderate and no response groups. Logistic regression analysis was applied to determine independent predictors of an MCII. Statistical analysis was done using Stata Inc software (v15) Stata Statistical Sofware: Release 15 (StataCorp. 2017. College Station, TX: StataCorp LCC.). A $P \leq 0.05$ was considered statistically significant.

Results

Of 5539 RA patients who had completed all 20 HAQ questions as of September 2014, 421 patients fulfilled the remaining inclusion criteria [Supplementary Figure 1]. Selected patients had the typical RA female predominance; long-standing established disease with very active disease [Table 1]. Over a mean standard deviation follow-up period of 8.7 (1.9) months, both disease activity and HAQ scores improved significantly, 199 (47%) of patients achieved a EULAR good/moderate response, of whom 64%, 63%, and 62% achieved an MCII with HAQ-8, HAQ-T, and HAQ-20 scoring, respectively.

The Bland–Altman plots [Supplementary Figure 2] for visit 1 and visit 2 showed good agreement between the three HAQ scoring methods. The agreement was better between HAQ-T and HAQ-20 than between HAQ-8 and other scoring methods. The effect size on the HAQ was larger in patients with good/moderate EULAR response than in patients with no EULAR response for all three HAQ scoring methods [Table 1].

Cumulative probability curves of changes in HAQ scores in relation to the two EULAR response groups [Figure 1], were similar for the three HAQ scoring methods. Of note, all three scoring methods showed that the proportion of patients who achieved an MCII was significantly higher in patients with good/moderate EULAR response than in patients with no EULAR response [Supplementary Table 1].

The strongest independent predictor of achieving a MCII in 269 (64%) patients (HAQ-8 method) was a EULAR good/moderate response (odds ratio [OR] = 7.11) [Table 2]. Other significant independent predictors of a MCII were high baseline disease activity score (OR = 1.55; 95% confidence interval [CI] = 1.20, 1.99), high baseline swollen

Table 1: Demographic, clinical characteristics, changes in disease activity, and Health Assessment Questionnaire scores in 421 rheumatoid arthritis patients

Variable			mean(SD) or *n*(%)		
Age (years)			55.0 (13.0)		
Females (%)			347/414 (84)		
Symptom duration (years)			10.5 (9.5)		
RF positive (%)			273/398 (68)		
ACPA positive (%)			146/357 (41)		
Duration between visits (months)			8.7 (1.9)		
Statistical test	**Visit 1**	**Visit 2**	***P***	**Effect size**	
				Guyatt	**Cohen's d**
DAS28-ESR*	6.1 (0.8)	4.8 (1.6)	<0.00001		
SJC*	8.1 (5.6)	4.9 (5.2)	<0.00001		
TJC*	12.9 (7.0)	8.0 (7.4)	<0.00001		
PGA (mm)*	67.6 (20.6)	52.3 (27.7)	<0.00001		
ESR (mm/h)*	44.2 (29.0)	35.6 (28.0)	<0.00001		
HAQ-8**	1.6 (1.1-2.1)	1.4 (0.9-1.9)	<0.00001	1.43 (1.2-1.6)	0.97 (0.8-1.2)‡
HAQ-T**	1.2 (0.7-1.6)	0.7 (0.4-1.4)	<0.00001	1.36 (1.1-1.6)	0.96 (0.8-1.2)‡
HAQ-20**	1.2 (0.8-1.6)	0.9 (0.5-1.4)	<0.00001	1.37 (1.1-1.6)	0.95 (0.7-1.0)‡

*Mean (SD), **Median (IQR), ‡Mean (95% CI). IQR: Interquartile range, CI: Confidence interval, RF: Rheumatoid factor, ACPA: Anti-citrullinated peptide antibody, ESR: Erythrocyte sedimentation rate, DAS28-ESR (4): Disease activity score with 28-joint count using the ESR (4 variable), PGA: Patient global assessment, SJC: Swollen joint count, TJC: Tender joint count, HAQ: Health Assessment Questionnaire, HAQ-8: HAQ-standard method, HAQ-T: HAQ-Tomlin method, HAQ-20: HAQ-20-item method, SD: Standard deviation

joint counts (OR = 1.04; 95% CI = 1.01,1.09) and tender joint counts (OR = 1.04; 95% CI = 1.01,1.07), and high baseline HAQ-8 score (OR = 1.06; 95% CI = 1.03,1.11).

Discussion

In this cohort of established, long-standing RA patients with high disease activity and in whom 47% attained a good/moderate EULAR response after a mean period of 8.7 months, the three methods showed little difference in sensitivity to change in disease activity. In the group of patients with a good/moderate response, 64% had a clinically meaningful improvement in functional disability, which was similar for the alternative HAQ scoring methods.

Our findings showed that there is good agreement between the three scoring methods and that effect size was large for all the methods demonstrating good but similar sensitivity to change in a patient group with high disease activity. That agreement was tighter between the two alternative scoring methods, HAQ-T and HAQ-20, compared to the traditional HAQ-8 is not surprising. Unlike the former methods, the HAQ-8 uses the worst scores per domain to score the HAQ. We were unable to confirm the assertions of Tomlin *et al*., which were based on results of a smaller sample size of 104 patients without accounting for disease activity, suggesting that the HAQ-T scoring method may achieve better sensitivity than HAQ-8 when used longitudinally.[6] Hence, the alternative scoring methods did not affect the overall sensitivity and therefore did not address the limitations of the floor and ceiling effects of the HAQ.[2]

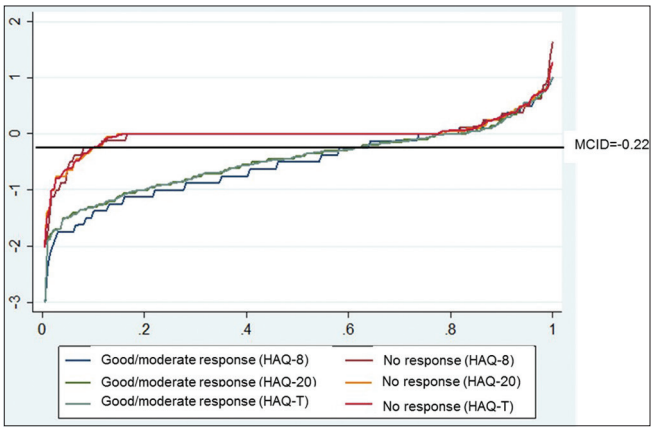


Figure 1: Cumulative probability curves of rheumatoid arthritis patients achieving a EULAR good/moderate response (*n* = 199) versus no response (*n* = 222) in relation to HAQ scores by three methods. EULAR: European League Against Rheumatism, HAQ: Health Assessment Questionnaire

Previous studies have shown that improvement in the functional disability is the greatest in early disease and declines with increasing disease duration.[17,18] While this is generally true, our findings underscore the importance of controlling disease activity. In the present study, significant improvements in functional disability were observed in patients who achieved a EULAR good/moderate response, despite having established long-standing disease, with a mean disease duration of just over 10 years. The depiction of change in HAQ scores in relation to EULAR RA response as cumulative probability curves is potentially a novel method to

Table 2: Analysis of baseline predictors of a minimally clinically important improvement by the HAQ-8 scoring method in 421 rheumatoid arthritis patients

Baseline variable	Achieved MCII (152), *n* (%)	Failed to achieve MCII (269), *n* (%)	Univariate analysis (*P*)	Multivariate analysis	
				OR (95% CI)	*P*
Age (years)*	53.7 (13.8)	56.0 (12.4)	0.091	-	-
Female gender	120 (81)	227 (86)	0.174	-	-
Symptom duration (months)*	7.8 (2.7)	7.9 (2.3)	0.701	-	-
RF positive	101 (72)	171 (74)	0.191	-	-
ACPA positive	58 (56)	88 (65)	0.164	-	-
Erosions present	61 (60)	131 (70)	0.078	-	-
Smoking, ever	77 (51)	144 (54)	0.571	-	-
DAS28-ESR (4)*	6.3 (0.8)	6.0 (0.8)	0.001	1.55 (1.20-1.99)	0.001
SJC (0-28)*	9.1 (6.0)	7.5 (5.3)	0.008	1.04 (1.01-1.09)	0.009
TJC (0-28)*	14.1 (7.0)	12.2 (6.9)	0.008	1.04 (1.01-1.07)	0.009
PGA VAS (mm)*	69.3 (20.4)	66.7 (20.6)	0.216	-	-
ESR (mm)*	45.3 (29.5)	43.5 (28.6)	0.540	-	-
Good/moderate EULAR response	127 (84)	72 (47)	<0.001	7.11 (3.67-13.80)	<0.001
HAQ-8*	1.3 (0.6)	1.1 (0.6)	0.001	1.06 (1.03-1.11)	0.001

*Mean (SD). OR: Odds ratio, CI: Confidence interval, RF: Rheumatoid factor, ACPA: Anti-citrullinated protein antibody, ESR: Erythrocyte sedimentation rate, DAS28-ESR (4): Disease activity score with 28-joint count using the ESR (4 variable), SJC: Swollen joint count, TJC: Tender joint count, PGA VAS: Patient global assessment on a visual analog scale, EULAR: European League Against Rheumatism, MCII: Minimally clinically important improvement, HAQ-8: HAQ-standard method, SD: Standard deviation

represent the changes in functional disability in clinical trials and longitudinal observational studies.[19,20]

Moreover, a EULAR good/moderate response, higher HAQ score, and high disease activity score, especially high swollen and tender joint counts, were independent predictors of achieving a HAQ MCII. In a multicenter prospective observational study of four European arthritis rehabilitation centers, Hagel *et al*. observed that health-related quality of life (HRQoL) improved most significantly in patients with the most severe symptoms.[21] Worse physical function as measured by the HAQ and more severe disease symptoms such as pain and fatigue at admission predicted an improved HRQoL postintervention. By contrast, other studies have shown higher baseline HAQ scores[22] and disease activity scores[23] to predict a poorer long-term functional outcome.

Limitations of the present study include the inability to account for comorbidities and surgical interventions as potential factors that impact on functional disability and the influence of specific classes of drugs such as biologics and corticosteroids. Also given that we preselected patients with high disease activity, improvements in disease activity and HAQ scores might represent a "regression to the mean" phenomenon.

Conclusion

Notwithstanding these limitations, our findings indicate no significant advantages of the different methods of scoring the HAQ in terms of sensitivity to change. Our results importantly show that even in patients with long-standing established disease, good control of disease activity is associated with clinically significant improvement in physical disability. The study underscores the importance of achieving EULAR good/moderate response, irrespective of disease duration.

Acknowledgments

The authors would like to thank Rosaline van den Berg and Sytske Anne Bergstra of the METEOR foundation for their guidance and advice, and Connective Tissue Disease Research Fund, University of the Witwatersrand for financial assistance.

Financial support and sponsorship

Nil.

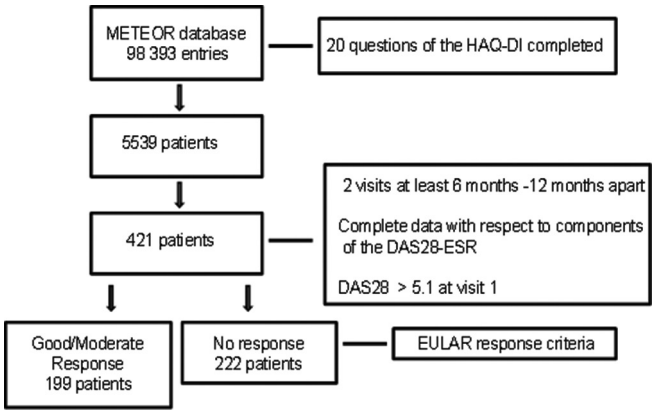
Conflicts of interest

There are no conflicts of interest.

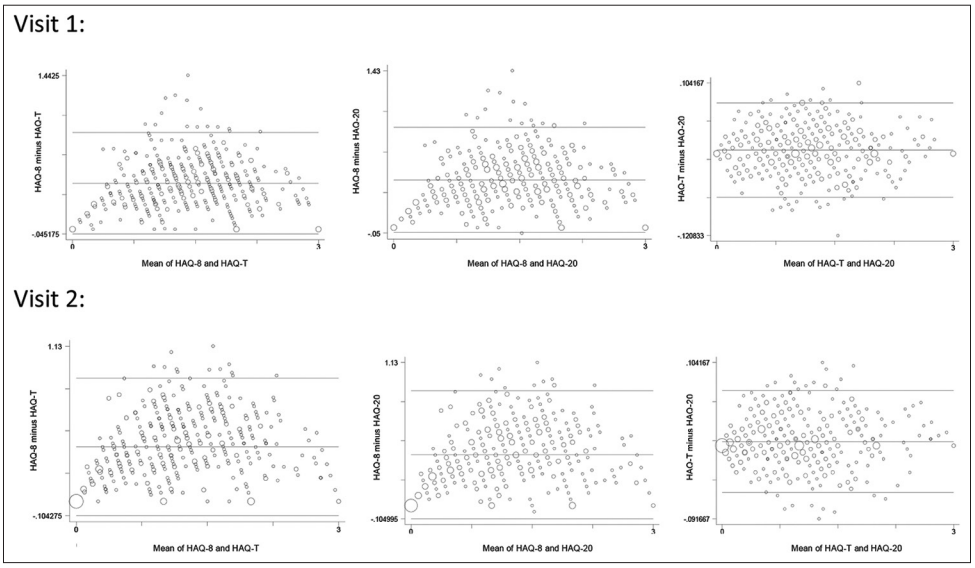
References

1. Barrett EM, Scott DG, Wiles NJ, Symmons DP. The impact of rheumatoid arthritis on employment status in the early years of disease: A UK community-based study. Rheumatology (Oxford) 2000;39:1403-9.
2. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: A review of its history, issues, progress, and documentation. J Rheumatol 2003;30:167-78.
3. Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components. Arthritis Rheum 2006;54:2784-92.
4. Aletaha D, Alasti F, Smolen JS. Chronicity of rheumatoid arthritis affects the responsiveness of physical function, but not of disease activity measures in rheumatoid arthritis clinical trials. Ann Rheum Dis 2015;74:532-7.
5. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.

6. Tomlin GS, Holm MB, Rogers JC, Kwoh CK. Comparison of standard and alternative health assessment questionnaire scoring procedures for documenting functional outcomes in patients with rheumatoid arthritis. J Rheumatol 1996;23:1524-30.
7. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): Analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. J Rheumatol 2001;28:982-9.
8. Bergstra SA, Machado PM, van den Berg R, Landewé RB, Huizinga TW. Ten years of METEOR (an International Rheumatoid Arthritis Registry): Development, research opportunities and future perspectives. Clin Exp Rheumatol 2016;34:S87-90.
9. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, *et al.* 2010 rheumatoid arthritis classification criteria: An American College of Rheumatology/European League against rheumatism collaborative initiative. Ann Rheum Dis 2010;69:1580-8.
10. 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.
11. van Riel PL, van Gestel AM, van de Putte LB. Development and validation of response criteria in rheumatoid arthritis: Steps towards an international consensus on prognostic markers. Br J Rheumatol 1996;35 Suppl 2:4-7.
12. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015;25:141-51.
13. van der Heijde D, Landewé R, Feldtkeller E. Proposal of a linear definition of the Bath Ankylosing Spondylitis Metrology Index (BASMI) and comparison with the 2-step and 10-step definitions. Ann Rheum Dis 2008;67:489-93.
14. Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171-8.
15. Sullivan GM, Feinn R. Using effect size–or why the *P* value is not enough. J Grad Med Educ 2012;4:279-82.
16. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: The patient's perspective. J Rheumatol 1993;20:557-60.
17. Aletaha D, Ward MM. Duration of rheumatoid arthritis influences the degree of functional improvement in clinical trials. Ann Rheum Dis 2006;65:227-33.
18. Aletaha D, Strand V, Smolen JS, Ward MM. Treatment-related improvement in physical function varies with duration of rheumatoid arthritis: A pooled analysis of clinical trial results. Ann Rheum Dis 2008;67:238-43.
19. Landewé R, van der Heijde D. Radiographic progression depicted by probability plots: Presenting data with optimal use of individual values. Arthritis Rheum 2004;50:699-706.
20. Kojima T, Takahashi N, Funahashi K, Asai S, Terabe K, Kaneko A, *et al.* Improved safety of biologic therapy for rheumatoid arthritis over the 8-year period since implementation in Japan: Long-term results from a multicenter observational cohort study. Clin Rheumatol 2016;35:863-71.
21. Hagel S, Lindqvist E, Petersson IF, Meesters JJ, Klokkerud M, Aanerud GJ, *et al.* Which patients improve the most after arthritis rehabilitation? A study of predictors in patients with inflammatory arthritis in Northern Europe, the STAR-ETIC collaboration. J Rehabil Med 2014;46:250-7.
22. Combe B, Cantagrel A, Goupille P, Bozonnat MC, Sibilia J, Eliaou JF, *et al.* Predictive factors of 5-year health assessment questionnaire disability in early rheumatoid arthritis. J Rheumatol 2003;30:2344-9.
23. Strating MM, Van Schuur WH, Suurmeijer TP. Predictors of functional disability in rheumatoid arthritis: Results from a 13-year prospective study. Disabil Rehabil 2007;29:805-15.



Supplementary Figure 1: Selection of rheumatoid arthritis patients with high disease activity from METEOR database. METEOR: Measurement of Efficacy of Treatment in the Era of Outcome in Rheumatology



Supplementary Figure 2: Bland–Altman plots illustrating agreement between three HAQ scoring methods in 421 rheumatoid arthritis patients. HAQ: Health Assessment Questionnaire

Supplementary Table 1: Minimally important clinical improvement by three Health Assessment Questionnaire scoring methods to European League Against Rheumatism responses in 421 rheumatoid arthritis patients

Scoring method	EULAR good/moderate response group (*n*=199), *n* (%)		EULAR no response group (*n*=222), *n* (%)		OR (95% CI)
	Achieved	Not achieved	Achieved	Not achieved	
HAQ-8	127 (64)	72 (36)	25 (11)	197 (89)	13.9 (8.4-23.1)
HAQ-T	125 (63)	74 (37)	24 (11)	198 (89)	13.9 (8.4-23.3)
HAQ-20	124 (62)	75 (38)	25 (11)	197 (89)	13.4 (8.1-22.2)

OR: Odds ratio, CI: Confidence interval, HAQ: Health Assessment Questionnaire, EULAR: European League Against Rheumatism, HAQ-8: HAQ-standard method, HAQ-T: HAQ-Tomlin method, HAQ-20: HAQ-20-item method