A Composite Disease Activity Scale for Clinical Practice, Observational Studies, and Clinical Trials: The Patient Activity Scale (PAS/PAS-II)

FREDERICK WOLFE, KALEB MICHAUD, and THEODORE PINCUS

ABSTRACT. ***Objective.*** To develop and validate a composite patient self-report disease activity scale for use in clinical practice and in observational studies and clinical trials.

Methods. A total of 9078 patients with rheumatoid arthritis completed detailed questionnaires that included measure of quality of life in the form of utilities. We evaluated several disease activity scales by measuring their agreement with the utility scales, and also their assessed ability to predict mortality and prescription for anti-tumor necrosis factor therapy.

Results. A composite index composed of a visual analog scale (VAS) for pain, a patient global VAS, and the Health Assessment Questionnaire (HAQ) or the HAQ II formed the Patient Activity Scale (PAS) and PAS-II. These scales performed as well as or better than longer, more complex scales.

Conclusion. A simple, useful clinical scale, the PAS or PAS-II, can be formed by the use of common clinical variables. It is well correlated with and relevant to a wide range of clinical variables. This scale should be useful for comparative studies, clinical care, and regulatory documentation. (J Rheumatol 2005;32:2410–5)

Key Indexing Terms:
PATIENT ACTIVITY SCALE
DISEASE ACTIVITY
HEALTH ASSESSMENT QUESTIONNAIRE
VISUAL ANALOG SCALE

Composite measures such as the Disease Activity Scale (DAS)[1-3] and the American College of Rheumatology (ACR) improvement criteria[4] are important single-score summary measures of improvement in clinical trials. The DAS also functions to provide an activity score, and can also be used in clinical practice. Recently, we have shown that a composite change score composed from the Health Assessment Questionnaire (HAQ)[5], a visual analog scale (VAS) for pain, and a VAS patient global severity scale, which are components of the ACR improvement criteria, performs as well as the full ACR criteria in 2 clinical trials[6].

Formal evaluation of patients with rheumatoid arthritis (RA) in the clinical practice setting is beset by numerous problems. Few North American rheumatologists formally evaluate patients with questionnaires or detailed joint examinations[7], and laboratory tests that are required for the DAS are usually not available on the day of the clinic visit. Overall, surveys have suggested that formal documented evaluation of RA disease activity is the exception in clinical practice.

Patient self-report data, if sufficiently comprehensive, could fill the need for a clinical summary score if certain conditions were met: a patient questionnaire should be simple to administer and score, provide easily interpretable results, allow comparison between individual patients and participants in clinical trials (and results of clinical trials), and be suitable for regulatory purposes. In addition, the scale should be suitable for use in observational studies.

Various composite indexes have been used over the last 50 years, but none has been based on patient self-report alone. However, the increasing recognition that patient self-report questionnaires may perform as well as standard physician and laboratory measures suggests the need for a patient composite index.

We combined HAQ, pain, and global severity scores to derive a Patient Activity Scale (PAS). In our validation of this scale, we address the question of how the components of the scale should be combined, and provide comparative data on how the scale compares with the widely used Medical Outcomes Study Short Form-36 (SF-36), as well as to the individual components and to other scales.

In the setting of a clinical trial, variable and composite scales can be evaluated for relative efficiency. In surveys and clinical practice, however, there is no natural "gold standard." In this report we use 2 quality of life utility scales as

From the National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine, Wichita, Kansas; Center for Primary Care and Outcomes Research, Stanford University, Stanford, California; and the Division of Rheumatology, Vanderbilt University, Nashville, Tennessee, USA.

F. Wolfe, MD, National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine; K. Michaud, MS, National Data Bank for Rheumatic Diseases and Stanford University; T. Pincus, MD, Division of Rheumatology, Vanderbilt University.

Address correspondence to Dr. F. Wolfe, National Data Bank for Rheumatic Diseases, Arthritis Research Center Foundation, 1035 N. Emporia, Suite 230, Wichita, KS 67214.
E-mail: fwolfe@arthritis-research.org

Accepted for publication July 25, 2005.

the standard and evaluate potential disease activity scales by the extent to which they agree with utility scores.

MATERIALS AND METHODS

Patient sample. Patients in this study were participants in the National Data Bank for Rheumatic Diseases (NDB) longitudinal study of RA outcomes. Patients are recruited from the practices of United States rheumatologists, and are followed with semiannual questionnaires[8-11]. This report deals with the status of 9078 patients with RA who completed study questionnaires between July 2002 and December 2004. For patients who completed more than one survey during this period a randomly selected survey was utilized.

Demographic and disease status variables. NDB participants completed semiannual, detailed 28 page questionnaires about all aspects of their illness. Demographic variables were recorded at each assessment including sex, age, ethnic origin, education level, current marital status, and medical history. Functional assessment measures included the Stanford Health Assessment Questionnaire (HAQ) functional disability index[5], the HAQ-II, a shortened, modified version with similar scaling but superior psychometric properties[12], and the SF-36, from which the physical component summary score (PCS) and mental component score (MCS) were calculated[13,14]. The Rheumatoid Arthritis Disease Activity Index (RADAI)[15,16] and a composite variable made up from the PCS and MCS were potential activity indexes that were compared against the PAS and PAS-II. We also collected data for the Regional Pain Scale (RPS), a measure of pain extent[8,17], and VAS scales for pain, fatigue, sleep disturbance and global severity[18]. Total medical costs were obtained as described[11].

To assess quality of life (QOL) by utilities, we administered the EuroQol[19-22], utilizing the newly developed US tariffs[23], the SF-6D[24], and a VAS based QOL scale utilizing linear transformation[25]. In our analyses, we used the QOL scales as a gold standard by which the PAS/PAS-II and other disease activity scales could be judged. As the SF-6D directly uses scales contained in the SF-36, we excluded that questionnaire from comparative analyses. The EuroQol contains 5 questions, 3 of which are about function, one about pain, and one about psychological status. The VAS QOL scale was anchored at one end with "death" and the other end with "perfect health." It was transformed using the algorithm[25]

$$0.44*(\text{VAS QOL}/100) + 0.49$$

VAS QOL scales have higher utilities than multi-item scales such as the EuroQol, and are closer in that respect to values obtained by the standard gamble[25]. Raw VAS scores are proportionally lower than true utility scores. Transformations alter the metric of the raw VAS scale so that its values and distribution closely represent values obtained by the standard gamble[25].

Statistical analyses. We formed the PAS and PAS-II by multiplying the HAQ (the HAQ-II for the PAS-II) by 3.33 and then dividing the sum of the VAS pain, VAS global, and HAQ/HAQ-II by 3. This yields a 0–10 scale. For test purposes we also made a composite SF-36 scale by taking the mean of the PCS and MCS.

We examined the strength of relationships between utility scores and disease activity scales using Kendall's tau-a. Tau-a has a simple interpretation, the percentage agreement between the utility scale and the disease activity scale. For example, a value of 0.59 (Table 1) indicates that it is 59% more likely that a person with a low utility value will have a high disease activity value than a low disease activity scale value. Higher utility values represent better health. Lower disease activity scores represent better health, except for the SF-36, in which scaling is reversed. Tau-a allows us to understand which disease activity measures are most strongly associated with QOL. In addition, the nonparametric tau-a is robust against extreme values produced by the EuroQol.

To evaluate whether the simple summation and averaging was a sufficient method of scale formation, we examined the first principal component of each scale by the Kendall tau-a comparison. In addition, we used fractional polynomial regression to determine if different weighting might yield a better scale. Kendall's tau-a and associated confidence limits were calculated using the Somers-D package[26]. For presentation in Table 1, negative values are presented as positive, for ease of reading the data.

Table 1. Association and comparison of PAS, PAS-II, and SF-36 combined score with EuroQol and VAS QOL utility scores (N = 9 078).

	Tau	z Score	p	95% CI
EuroQol				
PAS	0.59	136.5	0.000	0.59–0.60
PAS-II	0.59	135.1	0.000	0.58–0.60
SF-36 PCS + MCS	0.59	140.0	0.000	0.58–0.60
VAS QOL utility				
PAS	0.32	45.4	0.000	0.34–0.31
PAS-II	0.33	46.2	0.000	0.34–0.31
SF-36 PCS + MCS	0.35	51.1	0.000	0.34–0.37

Tau: Kendall's tau, PCS: physical component score, MCS: mental component score.

Cox proportional hazards regression was used to determine the effect of the various disease activity measures on (1) the risk of mortality and (2) the risk of a new prescription for an anti-tumor necrosis factor (TNF) treatment. For Cox analyses all patient observations between 2002 and mid-2004 were used, rather than a randomly selected observation as described above. The Bayesian information criterion (BIC)[27,28] is a goodness-of-fit measure of overall model fit and is a means to compare nested and non-nested models[29]. Based on the log-likelihood of the logistic regression, BIC values are useful to compare different models, but values have no directly interpretable meaning. Differences between models for the BIC of 0–2, 2–6, 6–10, and > 10 provide weak, positive, strong, and very strong evidence for the superiority of one model compared to another, according to Raferty[28]. BIC variable statistics can only be compared when sample sizes are the same[29]. All analyses were performed using Stata version 8.2[30].

RESULTS

Demographic and clinical characteristics of the 9078 study participants are presented in Table 2. The HAQ and HAQ-II scores are 1.08 and 1.04, respectively, indicating the close correspondence of the scales. This also results in PAS and PAS-II scores of 3.7 for both scales. The 3 utility measures each result in different utilities: EuroQol 0.73, VAS QOL 0.81, and SF-6D 0.63.

To compare the relative associations of composite variables, we used 2 utility measures as gold standards (Table 1), the EuroQol and the VAS QOL. The SF-6D was excluded because it shared items with the SF-36. We examined all clinical items in Table 2 for their ability to associate with the 2 utility scales. The PAS, PAS-II, and a composite score consisting of the mean of the SF-36 PCS and MCS performed far better than any other variables. These 3 variables were retained for further analyses. Using Kendall's tau as the measure of association, the PAS, PAS-II, and the SF-36 composite all showed equal strength of association with the EuroQol. When compared within the framework of the VAS QOL utility, the SF-36 performed marginally better, with a tau of 0.35 compared with 0.33.

When we assessed the ability of the RADAI, PCS, and MCS to associate with the EuroQol, values were RADAI 0.52 (95% CI 0.51–0.53), PCS 0.48 (95% CI 0.47–0.49),

Table 2. Demographic and clinical characteristics of 9 078 patients with RA.

Variable	Mean or %	SD
Age, yrs	62.2	12.6
Sex, % male	21.8	
Education category, yrs, %		
0–8	1.9	
8–11	6.4	
12	36.2	
13–15	26.4	
≥ 16	29.0	
Ethnic origin, %		
White, not of Hispanic origin	92.3	
Black, not of Hispanic origin	3.4	
Asian or Pacific Islander	0.9	
American Indian or Alaskan Native	1.0	
Hispanic	2.2	
Other	0.3	
Disease duration, yrs	16.2	10.9
HAQ (0–3)	1.08	0.7
HAQ II (0–3)	1.04	0.7
Pain (0–10)	3.8	2.7
Global severity (0–10)	3.7	2.4
Fatigue (0–10)	4.4	2.9
Physical component score, SF-36	32.0	10.4
Mental component score, SF-36	43.5	13.9
RADAI (0–10)	3.4	2.1
Regional pain scale (0–19)	5.7	5.0
EuroQol US (0–1)	0.73	0.2
VAS QOL (linear transformation)	0.81	0.1
SF-6D utility (0–1)	0.63	0.1
Patient activity score using HAQ (PAS) (0–10)	3.7	2.1
Patient activity score using HAQ II (PAS-II) (0–10)	3.7	2.1

and MCS 0.49 (95% CI 0.48–0.60). For the VAS QOL, the respective values were RADAI 0.27 (95% CI 0.26–0.29), PCS 0.31 (95% CI 0.29–0.32), and MCS 0.29 (95% CI 0.28–0.30). Therefore, the PAS and PAS-II were superior to the RADAI, PCS, and MCS for all of the utility scales evaluated, as shown in Table 1.

As it was possible that the combination variables were not weighted correctly, we extracted the first principal component of each index and compared its tau with the values in Table 1. The principal components offered improvements between 0.0 and 0.1 units. Similarly, we performed nonlinear regression to see if other methods of characterizing the composite variables improved their association with the utility scores. Only the slightest improvements were seen. Based on these results (not shown) and the advantages of simple scoring, we retained the PAS and PAS-II in their original forms. The distribution of the PAS scales and the SF-36 composite and principal component scales are shown in Figure 1.

We examined the performance of the new scales in predicting mortality, as shown in Table 3. We hypothesized that the HAQ and HAQ-II would be better predictors of mortality than the PAS scores, because the PAS functional component might be diluted by the pain and global severity components. One method of comparing models is through the use of the BIC, where lower scores may indicate a better model. Table 3 shows the PAS and PAS-II performing similarly, the HAQ-II performing better than the HAQ, and the SF-36 performing intermediate between the 2 HAQ. These

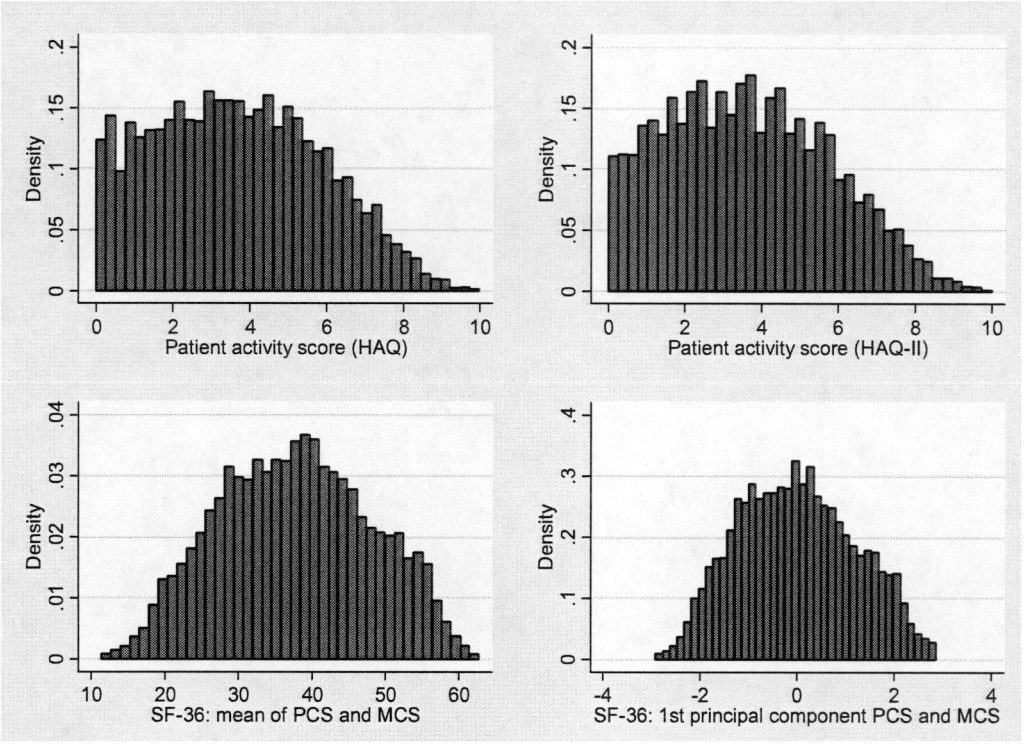

Figure 1. Histograms of PAS, PAS-II, SF-36 PCS/MCS composite, and first principal component of SF-36 PCS/MCS composite.

Table 3. Comparative predictive ability for mortality of PAS, PAS-II, HAQ, HAQ-II, and SF-36 combined score. Cox regression analyses of 12,433 patients during 26,806 patient-years of observation.

Predictor	Hazard Ratio	z Score	p	95% CI	BIC
PAS	1.3	11.2	0.000	1.3–1.4	6142.5
PAS-II	1.3	11.3	0.000	1.3–1.4	6141.1
HAQ	2.4	11.9	0.000	2.1–2.8	6121.7
HAQ-II	2.8	13.3	0.000	2.4–3.3	6091.0
SF-36 PCS + MCS	0.9	–12.6	0.000	0.9–0.9	6101.6

BIC: Bayesian information criterion, PCS: physical component score, MCS: mental component score.

data indicate satisfactory performance of the PAS scales as predictors of mortality, performing only marginally less well than the 2 HAQ.

We next used 3 types of scales to compare their ability to predict the institution of anti-TNF therapy in RA patients who had not yet received that therapy (Table 4). Based on the BIC, the PAS scales performed better than the HAQ scales, but not quite as well as the VAS pain scale. However, the differences between the PAS scales and pain scale were slight, indicating that the PAS scales satisfactorily predict anti-TNF use.

To place the scores of the PAS scales into the perspective of other clinical measures, we divided the PAS scales into quartiles and displayed the results of the other clinical scales for each quartile. As the results of the PAS and PAS-II were almost the same, we present only the PAS results in Table 5. With each quartile increase in the PAS, clinical variables worsen in a stepwise manner. This helps to place the PAS into a clinical perspective. Quartile I values indicate low disease activity. Quartile 4 values indicate very high disease activity (and increased association with mortality, as shown in Table 4). Quartiles 2 and 3 occupy intermediate positions.

Figure 2 shows a simple form that is suitable for rapid clinic use to calculate the PAS-II. Using the full HAQ would require an additional page.

DISCUSSION

We have shown that the Patient Activity Scales define levels of clinical activity, and that they adequately predict treatment change and mortality. The scale is simple to calculate and use in the clinic. In addition, as most randomized clinical trials (RCT) and observational studies collect data on pain, patient global, and HAQ, this allows the general use of the PAS without additional data collection. Among the benefits of this is to allow comparison among studies, and comparison of clinical patients to those who participate in studies, and to document clinical status for regulatory purposes.

Although we did not show how the PAS scales compare with standard measures such as the ACR criteria and the DAS in clinical trials, we have shown that the PAS components were effective predictors of response in a RCT[5]. We suspect that the PAS will be similarly effective, but it remains for RCT results to be analyzed to see if that is the case.

Among the benefits of the PAS scales is their simplicity, making them particularly useful in the clinic. The VAS and HAQ-II shown in Figure 2 allow the 3 scales to be scores

Table 4. Comparative predictive ability for starting and anti-TNF therapy for the first time for PAS, PAS-II, VAS pain, HAQ, and HAQ-II. Cox regression analyses of 4 005 patients during 6 943 patient-years of observation.

Predictor	Quartile	Hazard Ratio	z Score	p	95% CI	BIC
PAS	Q1	1.0				
	Q2	1.4	2.5	0.011	1.1–1.8	10585.7
	Q3	1.9	5.2	0.000	1.5–2.4	
	Q4	2.4	7.4	0.000	1.9–3.0	
PAS-II	Q1	1.0				
	Q2	1.4	2.3	0.021	1.0–1.8	10578.6
	Q3	2.0	5.8	0.000	1.6–2.6	
	Q4	2.4	7.4	0.000	1.9–3.1	
Pain	Q1	1.0				
	Q2	1.5	3.3	0.001	1.2–2.0	10571.9
	Q3	2.0	5.5	0.000	1.6–2.5	
	Q4	2.7	8.2	0.000	2.2–3.5	
HAQ	Q1	1.0				
	Q2	1.6	3.9	0.000	1.3–2.0	10598.9
	Q3	1.5	3.5	0.001	1.2–1.9	
	Q4	2.2	7.2	0.000	1.8–2.7	
HAQ-II	Q1	1.0				
	Q2	1.3	2.1	0.041	1.0–1.6	10620.0
	Q3	1.5	3.2	0.001	1.2–1.8	
	Q4	1.8	5.5	0.000	1.5–2.3	

BIC: Bayesian Information criterion.

Table 5. Mean scores for important clinical variables according to quartile of PAS/PAS-II.

Quartile	RADAI	PCS	MCS	RPS	Depression
1 (0.0–1.9)	1.3	43.5	54.7	2.1	1.4
2 (1.9–3.6)	2.6	33.7	46.8	4.3	2.0
3 (3.6–5.3)	4.0	28.2	40.6	6.6	2.8
4 (5.3–10)	5.9	22.6	31.6	9.9	3.9
Quartile	Fatigue	Sleep	Pain	HAQ	HAQ-II
1 (0.0–1.9)	1.7	1.6	1.0	0.29	0.35
2 (1.9–3.6)	3.6	3.0	2.6	0.89	0.85
3 (3.6–5.3)	5.2	4.3	4.6	1.31	1.22
4 (5.3–10)	6.9	5.8	7.2	1.85	1.75
Quartile	Global	VAS QOL	EuroQol	SF-6D	Costs ($)
1 (0.0–1.9)	1.1	0.86	0.88	0.73	6144
2 (1.9–3.6)	2.9	0.82	0.79	0.64	7033
3 (3.6–5.3)	4.4	0.79	0.72	0.59	7547
4 (5.3–10)	6.4	0.76	0.52	0.57	8563

RADAI: Rheumatoid Arthritis Disease Activity Index; PCS: SF-36 physical component score; MCS: SF-36 mental component score; RPS: regional pain scale; Depression: Arthritis Impact Measurement Scales depression score; Fatigue: VAS fatigue score; Sleep: VAS sleep disturbance score; Pain: VAS pain score; HAQ: Health Assessment Questionnaire Disability Index; HAQ-II: Health Assessment Questionnaire Disability Index-II; Global: patient global severity; VAS QOL: linearly transformed VAS QOL utility scale; EuroQol: EuroQol utility score; SF-6D: SF-6D utility score; Costs: total direct medical costs (2001 dollars).

Considering ALL THE WAYS THAT YOUR ILLNESS AFFECTS YOU, RATE HOW YOU ARE DOING on the following scale. Place an X in the box below that best describes how you are doing on a scale of 0-10.

VERY WELL **0** ☐ **10** VERY POOR

We are interested in learning how your illness affects your ability to function in daily life. Place an X in the box which best describes your usual abilities OVER THE PAST WEEK:

Are you able to:	**Without Any Difficulty (0)**	**With Some Difficulty (1)**	**With Much Difficulty (2)**	**Unable To Do (3)**
Stand up from a straight chair?	☐	☐	☐	☐
Walk outdoors on flat ground?	☐	☐	☐	☐
Get on/off toilet?	☐	☐	☐	☐
Reach and get down a 5 pound object (such as a bag of sugar) from just above your head?	☐	☐	☐	☐
Open car doors?	☐	☐	☐	☐
Do yard work (outside work or activities)?	☐	☐	☐	☐
Wait in a line for 15 minutes?	☐	☐	☐	☐
Lift heavy objects?	☐	☐	☐	☐
Move heavy objects?	☐	☐	☐	☐
Go up two or more flights of stairs?	☐	☐	☐	☐

We are interested in knowing about any problems that you may have been having with fatigue. How much of a problem has fatigue or tiredness been for you IN THE PAST WEEK? Place an X in the box below that best describes the severity of your fatigue on a scale of 0-10.

FATIGUE IS NO PROBLEM **0** ☐ **10** FATIGUE IS A MAJOR PROBLEM

We are also interested in learning whether or not you are affected by pain because of your illness.

How much pain have you had because of your illness in the past week? Place an X in the box that best describes the severity of your pain on a scale of 0-10.

NO PAIN **0** ☐ **10** SEVERE PAIN

Figure 2. One-page questionnaire suitable for clinic use and PAS-II scoring.

merely by looking at them. To compute the PAS, the HAQ-II is multiplied by 3.3 and the sum of the 3 components is divided by 3. This process should take under 15 seconds.

It is of interest that the simple PAS scale performed just about as well as the SF-36 composite scale. We used the composite scale to maximize the effect of the SF-36 for study purposes. The SF-36 PCS and MCS scales alone were inferior to the PAS, as was the RADAI. In addition, the SF-36 requires a computer for scoring. Our results also showed that complex methods of weighting scale components were not required.

In summary, we have shown that a simple, useful clinical scale, the PAS and PAS-II, can be formed by the use of common clinical variables, and demonstrated its relevance to clinical variables. These scales should be useful for comparative studies, clinical care, and regulatory documentation. The PAS-II scale, as it is based on the shorter HAQ-II, may be preferred by clinical rheumatologists for its ease of use.

REFERENCES

1. van der Heijde DMFM, van't Hof MA, van Riel PLCM, van Leeuwen MA, van Rijswijk MH, van de Putte LBA. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Ann Rheum Dis 1992;51:177-81.
2. den Broeder AA, Creemers MC, van Gestel AM, van Riel PL. Dose titration using the Disease Activity Score (DAS28) in rheumatoid arthritis patients treated with anti-TNF-alpha. Rheumatology Oxford 2002;41:638-42.
3. van Gestel AM, Anderson JJ, van Riel PL, et al. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. J Rheumatol 1999;26:705-11.
4. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727-35.
5. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
6. Pincus T, Strand V, Koch G, et al. An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from that of placebo as effectively as the American College of Rheumatology 20% response criteria or the Disease Activity Score in a rheumatoid arthritis clinical trial. Arthritis Rheum 2003;48:625-30.
7. Wolfe F, Pincus T, Thompson AK, Doyle J. The assessment of rheumatoid arthritis and the acceptability of self-report questionnaires in clinical practice. Arthritis Rheum 2003;49:59-63.
8. Wolfe F, Michaud K. Severe rheumatoid arthritis, worse outcomes, comorbid illness, and sociodemographic disadvantage characterize rheumatoid arthritis patients with fibromyalgia. J Rheumatol 2004;31:695-700.
9. Wolfe F, Michaud K, Anderson J, Urbansky K. Tuberculosis infection in patients with rheumatoid arthritis and the effect of infliximab therapy. Arthritis Rheum 2004;50:372-9.
10. Wolfe F, Michaud K. Heart failure in rheumatoid arthritis: rates, predictors, and the effect of anti-tumor necrosis factor therapy. Am J Med 2004;116:305-11.
11. Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in persons with rheumatoid arthritis: a 3 year study of 7,527 patients. Arthritis Rheum 2003;48:2750-62.
12. Wolfe F, Michaud K, Pincus T. HAQ-II: Development and validation of a revised version of the Health Assessment Questionnaire (HAQ). Arthritis Rheum 2004;50:3296-305.
13. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40-66.
14. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
15. Stucki G, Liang MH, Stucki S, Bruhlmann P, Michel BA. A self-administered Rheumatoid Arthritis Disease Activity Index (RADAI) for epidemiologic research. Psychometric properties and correlation with parameters of disease activity. Arthritis Rheum 1995;38:795-8.
16. Fransen J, Langenegger T, Michel BA, Stucki G. Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. Rheumatology Oxford 2000;39:321-7.
17. Wolfe F. Pain extent and diagnosis: development and validation of the regional pain scale in 12,799 patients with rheumatic disease. J Rheumatol 2003;30:369-78.
18. Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. J Rheumatol 2004;31:1896-902.
19. Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. Ann Rheum Dis 2004;63:723-9.
20. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997;36:551-9.
21. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. Br J Rheumatol 1997;36:786-93.
22. EuroQol Group. EuroQol — a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
23. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203-20.
24. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271-92.
25. Mrus JM, Yi MS, Freedberg KA, et al. Utilities derived from visual analog scale scores in patients with HIV/AIDS. Med Decis Making 2003;23:414-21.
26. Newson R. Parameters beyond "nonparametric" statistics: Kendall's tau, Somer's D and median differences. Stata Journal 2002;2:45-64.
27. Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6:461-4.
28. Raferty AE. Bayesian model selection in social research. In: Marsden PV, editor. Sociological methodology. Oxford: Basil Blackwell; 1996:111-63.
29. Long JS, Freeze J. Regression models for categorical dependent variables using Stata. College Station, TX: Stata Press; 2001.
30. Stata Corporation. Stata statistical software: Release 8.2. College Station, TX: Stata Corporation; 2003.