Development and Validation of a Preference Weight Multiattribute Health Outcome Measure for Rheumatoid Arthritis

CHIUN-FANG CHIOU, MARIA E. SUAREZ-ALMAZOR, CATHY D. SHERBOURNE, CHIH-HUNG CHANG, CAROLINA REYES, MICHELLE DYLAN, JOSHUA OFMAN, DANIEL J. WALLACE, WESLEY MIZUTANI, and MICHAEL WEISMAN

ABSTRACT. ***Objective.*** To develop and validate multiattribute measures for patients with rheumatoid arthritis (RA) to report health states and estimate preference weights.

Methods. Survey materials were mailed to 748 patients. Factor analysis, an item response theory-based model, and an internal consistency test were used to identify attributes and evaluate items. Two multiattribute preference weight functions (MAPWF) were constructed. Construct validity of the new measures was then tested.

Results. Four hundred eighty-seven patients returned the survey; 24 items on 6 health attributes were selected to form the new outcomes measure. Two MAPWF were derived with preference weights measured with time tradeoff and visual analog scales as dependent variables. All validity test results were statistically significant.

Conclusion. Our results reveal that the new measures are reliable and valid in assessing health states and associated preference weights of patients with RA. (First Release Oct 15 2006; J Rheumatol 2006;33:2409–11)

Key Indexing Terms:
RHEUMATOID ARTHRITIS PREFERENCE WEIGHT DEVELOPMENT VALIDATION

With ever-increasing demand for health services from patients and constrained budgets, healthcare decision-makers are interested in the relative effectiveness and cost-effectiveness of medical interventions for a given disease as well as across various diseases. To compare effectiveness among various medical interventions, the outcome measure Quality-Adjusted life years (QALY) has been commonly used and recommended[1]. To calculate QALY, it is usually difficult for researchers to find a set of preference weights since generic health-related quality of life (HRQOL) measures[2-7] may not be sensitive enough to detect changes in health states, and direct preference elicitation techniques [i.e., visual analog scales (VAS), time tradeoff (TTO), and standard gamble] utilize numerous classifications of health outcomes. Our objective was to develop and validate multiattribute measures that can be used by patients with rheumatoid arthritis (RA) when reporting their health states, and to estimate preference weights for these health states.

From Amgen, Inc., Thousand Oaks, California; Health Services Research, Baylor College of Medicine, Houston, Texas; Health Program, RAND Corporation, Santa Monica, California; Feinberg School of Medicine, Northwestern University, Chicago, Illinois; Genentech, South San Francisco; Cerner LifeSciences, a Cerner Company, Beverly Hills; Departments of Medicine and Rheumatology, Cedars-Sinai Health System, Los Angeles; and Talbert Medical Group, Huntington Beach, California, USA.

Supported by a grant from Amgen, Inc.

C-F. Chiou, PhD, Amgen, Inc; M.E. Suarez-Almazor, MD, PhD, Health Services Research, Baylor College of Medicine; C.D. Sherbourne, PhD, Health Program, RAND Corporation; C-H. Chang, PhD, Feinberg School of Medicine, Northwestern University; C. Reyes, PhD, Genentech; M. Dylan, PhD, Senior Research Associate, Cerner LifeSciences; J. Ofman, MD, MSHS, Amgen, Inc; D.J. Wallace, MD, Departments of Medicine and Rheumatology, Cedars-Sinai Health System; W. Mizutani, MD, Talbert Medical Group; M. Weisman, MD, Departments of Medicine and Rheumatology, Cedars-Sinai Health System.

Address reprint requests to Dr. M. Dylan, Cerner LifeSciences, 9100 Wilshire Blvd., Suite 655E, Beverly Hills, CA 90212, USA. E-mail: mdylan@cerner.com

Accepted for publication August 2, 2006.

MATERIALS AND METHODS

A multistep process was undertaken to (1) develop the multiattribute questionnaire to enable patients to self-report their health states; (2) generate an algorithm to estimate preference weights; and (3) examine the construct validity and internal consistency of the new measures.

Step 1. An item pool derived from previous HRQOL instruments, patient focus groups, and experts' input[8] was reviewed by the study steering committee. The committee selected 28 items through consensus for the initial questionnaire (Table 1).

Questionnaire responses were used for principal component factor analysis with oblique promax rotation, reliability analysis with Cronbach's alpha coefficient, and Rasch analysis with Andrich's rating scale model[9-11]. Ceiling and floor effects were also examined.

The questionnaire, along with VAS health states preference questions, TTO general health preference question, Mannequin Forms for joint count evaluations[12], Stanford Health Assessment Questionnaire (HAQ) Disability Index[13], and Medical Outcomes Study short-form 36 (SF-36) questionnaire[5], was mailed to 748 patients with RA from Southern California. Institutional review board approval for the study was obtained from Cedars-Sinai Medical Center.

Step 2. The multiattribute preference weight function (MAPWF) technique[2,6,7] was adapted to develop an algorithm to estimate preference weights for all health states measured by the new questionnaire. The dependent MAPWF variable was the preference weight for the current health state measured with different techniques. Independent variables were the mean of items

Table 1. The 28 items for developing a questionnaire for measuring health states of patients with RA.

Question	
During the past 2 weeks, how difficult was it for you to	[Scale Not at all difficult (3); Somewhat difficult (2); Very difficult (1); Extremely difficult (0)]
1. Do such things as dress yourself or wash and dry yourself? (PHYS1)	
2. Use a knife or fork? (PHYS2)	
3. Open a jar or turn a doorknob? (PHYS3)	
4. Get in and out of bed? ((PHYS4)	
5. Climb one flight of stairs? (PHYS5)	
6. Walk one block? (PHYS6)	
7. Do paid work, household work, or school work? (PHYS7)	
8. Participate in leisure activities outside of your home, such as going out to dinner or movies or visiting neighbors or friends? (PHYS8)	
During the past 2 weeks, how often . . .	[Scale: Never (3); Sometimes (2); Often (1); Always (0)]
9. Were you worried about your rheumatoid arthritis (RA)? (EMOT1)	
10. Did you feel downhearted and blue? ((EMOT2)	
11. Did you worry about the side effects of your medication? (EMOT3)	
12. Did you feel calm and peaceful? (EMOT4)	
13. Were you frustrated because you didn't get relief from your RA? (EMOT5)	
14. Were you concerned about the impact of RA on your physical appearance? (EMOT6)	
15. Did you wake up at night because of pain or discomfort? (SYMP1)	
16. Did pain limit you in your daily activities? (SYMP2)	
17. Did you feel tired because of your RA? (SYMP3)	
18. Did fatigue limit your daily activities? (SYMP4)	
19. Did you feel full of energy? (SYMP5)	
20. Were you bothered by side effects from your medication? (THER1)	
21. Did side effects from your medication limit your daily activity? (THER2)	
22. Did you spend time with others? (SOCI1)	
23. Were you on the phone with friends or relatives? (SOCI2)	
24. Did you feel supported by your family or friends? (SOCI3)	
25. Did your family and friends express an interest in helping you with your problems? (SOCI4)	
During the past 2 weeks, how . . .	[Scale: Not severe at all (3); Somewhat severe (2); Very severe (1); Extremely severe (0)]
26. Severe were the side effects from your medication? (SYMP6)	
27. Severe was your overall pain? (THER3)	
28. Comfortable were you in asking for help when needed? (SOCI5)	

of each attribute, interaction of any 2 of these means, and demographic factors including age, sex, ethnicity, working status, and education.

Assuming an additive functional form of the MAPWF, ordinary least-squares stepwise regression, including first-order interactions, was employed. Only variables with a $p < 0.20$ were retained.

Step 3. Cronbach's alpha coefficient was calculated for the internal consistency for each attribute (0.70 or higher indicates good internal consistency[14]). Tests for convergent validity included calculating Pearson correlation coefficients between the estimated preference weights and SF-36 Mental Component Summary (MCS) and Physical Component Summary (PCS), pain and tender joint count, swollen joint count, and HAQ Disability Index.

To determine whether the new measures can estimate differences between patients or changes in patient status over time, survey respondents were divided into 6 groups of differing severity based on their HAQ scores (lowest severity if HAQ score 0 to 0.5; highest severity if HAQ score 2.5 to 3). Mean estimated preference weights for these 6 groups were compared using analysis of variance (ANOVA) and Kruskal-Wallis test.

RESULTS

Among the 748 patients who agreed to participate, 487 returned the survey (response rate 65.1%). The sample consisted of 70% Caucasians, 78.8% women, and 49.4% with at least a 2-year college degree. The average SF-36 MCS and PCS scores of these patients were 36.6 ± 9.4 and 43.7 ± 10.8 mean ± SD, respectively. The mean HAQ score was 1.13 ± 0.75.

Factor analysis revealed that 6 factors should be retained: Factor 1: physical function with items 4 to 8; Factor 2: RA symptom distress with items 12, 13, 15-19, and 26; Factor 3: social interaction with items 22-25 and 28; Factor 4: treatment-related symptom distress with items 20, 21, and 27; Factor 5: dexterity with items 1-3; and Factor 6: emotion with items 9-11 and 14. Results of Rasch analysis and Cronbach's alpha coefficient test informed the removal of items 4, 12, 19, and 28 from the questionnaire. Internal consistency was improved after item reduction and was greater than 0.70. No item had either floor or ceiling effects.

The first MAPWF with TTO as the dependent variable resulted in physical function, emotion, RA symptom distress, dexterity, interaction of physical function and RA symptom dis-

Table 2. Algorithm for predicting preference weights measured with time tradeoff (TTO) and visual analog scale (VAS) from the multiattribute preference weight functions (MAPWF).

Variable	VAS		TTO	
	Coefficient	p	Coefficient	p
Intercept	0.0369	< 0.0001	0.2985	< 0.0001
Physical	–0.0185	0.4784	0.0543	0.1409
Emotion	0.0051	0.8165	0.0927	0.1015
RA symptom	0.1006	0.2527	–0.0544	0.4001
Therapy	0.0711	0.1428		
Dexterity	0.0747	0.1180	0.0773	0.0041
Physical × RA symptom	0.0559	0.0119	0.0433	0.1009
Emotion × therapy	–0.0475	0.0882		
RA symptom × therapy	0.0452	0.1218		
RA symptom × dexterity	–0.0719	0.0378		
Physical × emotion			–0.0461	0.0788
Emotion × dexterity	0.0476	0.0942		
Female	–0.0403	0.1154	0.0644	0.0354
College			0.0463	0.1241
White	–0.0413	0.0363		

tress, interaction of physical function and emotion, sex, and marital status being the independent variables retained for the algorithm. The second MAPWF used VAS as the dependent variable and resulted in the following independent variables being retained: physical function, emotion, RA symptom distress, therapy-related distress, dexterity, interaction of physical function and RA symptom distress, interaction of emotion and therapy-related distress, interaction of RA symptom distress and therapy-related distress, interaction of RA symptom distress and dexterity, interaction of emotion and dexterity, sex, and race (Table 2). Both VAS and TTO were considered as direct preference elicitation techniques since they are the most commonly used.

Convergent validity was supported by the Pearson correlation coefficients (all $p < 0.0001$), which ranged from –0.3913 (correlation between predicted TTO preference weight and swollen joint count) to 0.7801 (correlation between predicted VAS preference weight and SF-36 PCS score). With 6 severity levels based on their HAQ scores, the mean predicted preference weights ranged from 0.8222 for patients with the least severe RA to 0.2907 for those with the most severe disease. Both ANOVA and Kruskal-Wallis tests indicated that the new measure has good discriminant validity ($p < 0.0001$).

DISCUSSION

Our findings identified and validated 24 items for 2 new outcomes measures for patients with RA that assess the impact of RA on 6 health attributes. The new measures can serve as HRQOL instruments for patients with RA. Further, the algorithms attached to the measures can estimate the preference weights for health states and thus enable researchers to estimate and compare QALY across different interventions and from different studies.

ACKNOWLEDGMENT

We thank the staff of Dr. Michael Weisman for their assistance throughout the study.

REFERENCES

1. Gold MR, Patrick DL, Torrance GW, et al. Identifying and valuing outcomes. In: Gold M, Siegal J, Russel L, Weinstein M, editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996:82-134.
2. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. Health Utilities Index. Pharmacoeconomics 1995;7:503-20.
3. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. Med Care 1996;34:702-22.
4. EuroQol — a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16:199-208.
5. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
6. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. J Clin Epidemiol 1998;51:1115-28.
7. Nichol MB, Sengupta N, Globe DR. Evaluating quality-adjusted life years: estimation of the health utility index (HUI2) from the SF-36. Med Decis Making 2001;21:105-12.
8. Weisman MH, Paulus HE, Russak SM, et al. Development of a new instrument for rheumatoid arthritis: the Cedars-Sinai Health-Related Quality of Life instrument (CSHQ-RA). Arthritis Rheum 2003;49:78-84.
9. Andrich D. A binomial latent trait model for the study of Likert-style attitude questionnaires. Br J Math Stat Psychol 1978;31:84-98.
10. Andrich D. Scaling attitude items constructed and scored in the Likert tradition. Educ Psychol Meas 1978;38:665-80.
11. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Nielson and Lydiche; 1960.
12. Wong AL, Wong WK, Harker J, et al. Patient self-report tender and swollen joint counts in early rheumatoid arthritis. Western Consortium of Practicing Rheumatologists. J Rheumatol 1999;26:2551-61.
13. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
14. Scientific Advisory Committee. Instrument review criteria. Medical Outcomes Trust Bulletin I-IV 1995. Waltham, MA: Medical Outcomes Trust; 1995.