

Adaptation and Validation of the Rheumatoid Arthritis Quality of Life Scale for Use in Canada

CAROLYN NEVILLE, DIANE WHALLEY, STEPHEN McKENNA, MARTINE Le COMTE, and PAUL R. FORTIN

ABSTRACT. Objective. The Rheumatoid Arthritis Quality of Life questionnaire (RAQoL) was developed simultaneously in the UK and the Netherlands to measure quality of life in patients with RA. We adapted and validated the RAQoL for the English-Canadian and French-Canadian languages and culture.

Methods. The UK RAQoL was translated into French-Canadian by a bilingual translation panel. Separate lay panels were then used to ensure that this and the English-Canadian instruments were appropriate for use with Canadian patients. Interviews were conducted with 15 French-Canadian and 15 English-Canadian patients with RA to determine the content validity. Reliability and construct validity were established by means of test-retest mail surveys conducted with 92 French-Canadian and 87 English-Canadian RA patients. The survey consisted of the adapted RAQoL, the Health Assessment Questionnaire (HAQ), and a demographic questionnaire.

Results. The RAQoL was successfully adapted for both the French and English-Canadian cultures. Field testing showed both versions to be well received by respondents. Of the French-Canadian patients included in the postal survey, 52 responded at Time 1 and 50 at Time 2. For the English-Canadian sample, 54 responded at both time points. Missing data rates for the RAQoL were low and floor and ceiling effects were minimal. Test-retest reliability was good for both versions: 0.87 for the French-Canadian and 0.95 for the English-Canadian. Alpha coefficients (0.92 for the French-Canadian, 0.93 for the English-Canadian) showed the items to be adequately interrelated and scores on the measure showed moderate to high correlations with the HAQ, confirming construct validity. Both versions of the RAQoL were also able to distinguish patient groups that differed according to perceived health status and perceived severity of RA. In addition, the French-Canadian version was able to distinguish patients who rated today as bad or very bad from those who rated today as good or very good.

Conclusion. The new versions of the RAQoL were well received by both French and English speaking Canadians. The psychometric quality of the adapted questionnaires means they are suitable for inclusion in clinical trials involving patients with RA. (*J Rheumatol* 2001;28:1505-10)

Key Indexing Terms:

RHEUMATOID ARTHRITIS
HEALTH STATUS

ADAPTATION

QUALITY OF LIFE
VALIDITY

The Rheumatoid Arthritis Quality of Life (RAQoL) questionnaire¹ was developed simultaneously in the United Kingdom and The Netherlands to measure quality of life in patients with rheumatoid arthritis (RA). It is a disease-

specific instrument that has the needs-based model of QoL as its theoretical basis. The content of the measure was derived from interviews with patients with RA, with the actual words of the patients used where possible. Such methodology was employed to maximize the immediacy of the resulting instrument to future respondents. The RAQoL consists of 30 statements describing relevant experiences in the lives of patients with RA. Each statement has a simple yes/no response format. Field testing in the UK and the Netherlands showed the measure to be a practical tool, taking about 6 minutes to complete. It is highly relevant and acceptable to patients and is easy to administer and score. Both UK English and Dutch language versions have been shown to have high internal consistency and test-retest reliability and to have good construct validity².

The RAQoL is an RA-specific QoL measure that has the potential to be of value in the measurement of outcome in clinical trials or longitudinal observational studies. The instrument serves as a useful complement to existing

From the Division of Clinical Epidemiology, McGill University Health Centre, Montreal, Quebec, Canada; Galen Research, Manchester, UK; and the Arthritis Centre of Excellence, Toronto Western Hospital, Toronto, Ontario, Canada.

Supported by Novartis Pharma Canada Inc. and Novartis Pharma AG, Basel, Switzerland. Dr. Fortin is a Senior Research Scholar of The Arthritis Society.

C. Neville, RN, BA; M. Le Comte, DEC, Division of Clinical Epidemiology, McGill University Health Centre; D. Whalley, BSc; S.P. McKenna, PhD, Galen Research; P.R. Fortin, MD, MPH, FRCPC, Director of Clinical Research, Arthritis Centre of Excellence, Toronto Western Hospital, University Health Network, and Associate Professor of Medicine, University of Toronto.

Address reprint requests to Dr. P.R. Fortin, Arthritis Centre of Excellence, Toronto Western Hospital, Room FP-1-214, 399 Bathurst Street, Toronto, Ontario, Canada M5T 2S8. E-mail: paul.fortin@uhn.on.ca

Submitted August 22, 2000; revision accepted January 23, 2001.

outcome instruments, particularly the Stanford Health Assessment Questionnaire (HAQ)³. The HAQ measures physical disability as determined by the patient's ability to perform certain physical tasks and activities. The RAQoL assesses QoL in terms of needs that are being interfered with as a result of having RA and the associated physical disability. When used together, the instruments provide a more complete picture of the impact of RA. Our purposes were (1) to adapt the RAQoL into the French-Canadian and English-Canadian languages and cultures, and (2) to validate the resulting new versions of the instrument.

MATERIALS AND METHODS

The initial stages of the production of the French-Canadian RAQoL adopted the dual panel translation method⁴. This approach emphasizes conceptual rather than linguistic translation and involves 2 translation panels, a bilingual and a lay panel. Production of the English-Canadian version followed a similar methodology, but the initial bilingual translation panel was not required. The subsequent testing of the 2 new versions followed the methodology employed in the original testing of the RAQoL². The specific stages in the adaptation were as follows.

Bilingual translation panel. The panel consisted of 5 persons who were fluent in French-Canadian and English. The meeting was also attended by one of the original authors of the RAQoL, whose role was to describe the meaning of the UK English items to the panel. The purpose of the panel was to produce the initial French-Canadian translation of the questionnaire. Emphasis was placed on obtaining translations for the items and instructions that were conceptually equivalent to those of the original UK measure. The translations were also required to be phrased in a way that would be acceptable to future Canadian respondents.

The questionnaire was first displayed on a screen to allow the panel members to visualize the measure as a whole. Each line of instruction and each question from the RAQoL was then discussed individually. Panel members were encouraged to agree on a single translation. However, where a consensus could not be reached, alternative translations were produced for consideration by the lay panel.

Lay panel. A separate lay panel was conducted for each language version. The French-speaking panel considered the initial translation produced by the bilingual panel, whereas the English-speaking panel worked with the original UK RAQoL. The purpose of both lay panels was to ensure that the new versions of the RAQoL were appropriate for completion by the average Canadian patient. They were allowed to modify the wording of the items and instructions if necessary as long as the original meaning was not altered. Each group comprised 6 persons of average or below average education level who were recruited from outside the hospital and who were not health care providers. The English-speaking panel was also attended by one of the original authors of the RAQoL to clarify the meaning of the UK items if required. The format of the groups was similar to that employed in the bilingual panel. The groups discussed the appropriateness of the wording to convey the intended meaning.

Field test interviews. The purpose of these interviews was to test the face and content validity of the new versions of the questionnaire⁵. Thirty patients with RA were recruited from the rheumatology clinic at the Montreal General Hospital and from 2 private rheumatology practices. The French-Canadian version of the RAQoL was tested with 15 French-Canadian speakers and the English-Canadian version was tested with 15 English-Canadian speakers. The interviews were conducted on a one-to-one basis in a private office in the Department of Rheumatology of the Montreal General Hospital.

Patients were asked to complete the RAQoL and to comment on its clarity and the appropriateness of its content. The amount of time needed to complete the measure was also recorded.

Postal surveys. The purpose of this stage was to test the reliability and construct validity of the new versions of the RAQoL. Ninety-two French-Canadian and 87 English-Canadian patients with RA were identified at the rheumatology clinics of the Montreal General Hospital and Hôpital Maisonneuve-Rosemont. Patients were not asked in advance to participate. The postal packages included the relevant version of the RAQoL, the HAQ, a 2 page demographic questionnaire that included patients' assessments of their own disease severity and disease activity, a letter of introduction explaining the survey, a self-addressed stamped envelope, and a refusal card to give the patients an opportunity to decline to participate.

After 2 weeks, patients who had not responded to the mailing in any way were telephoned to determine their willingness to participate.

Respondents were sent a second identical package 2 weeks after the first administration and again were reminded by telephone contact if the second mailing was not received within the 2 week time period.

Copies of the original RAQoL may be obtained from Diane Whalley, Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North, Manchester, UK, M15 6SE. The French-Canadian and English-Canadian versions of the RAQoL may be obtained from Dr. P. Fortin, Toronto Western Hospital, 399 Bathurst Street, Room FP-1-214, Toronto, Ontario, Canada M5T 2S8.

Statistical analysis. The returned questionnaires were photocopied and mailed to Galen Research, UK, for data entry and analysis. Nonparametric statistical analyses were applied throughout, as none of the instruments produced interval level data. Statistical significance was at the 95% level. Each language version was treated separately.

Test-retest reliability represents the reproducibility of the measure over time, assuming that no change has occurred between administrations. It was assessed by correlating RAQoL scores (using Spearman rank correlation coefficients) on the 2 administrations. A high correlation (> 0.85)⁶ indicates that the instrument is reliable, that is, that it produces little random measurement error.

Internal consistency was assessed using Cronbach's alpha coefficients. Alpha measures the extent to which the items in a scale are interrelated. A low alpha (< 0.70)⁷ is indicative of individual items not contributing adequately to the overall scale.

Construct validity is the extent to which an instrument measures the intended construct. This was estimated by determining convergent validity and known-groups validity. Convergent validity concerns the degree of association with comparator scales that measure a related construct. The HAQ was used as the comparator for both language versions of the RAQoL. A moderate to high correlation was expected between the measures. Known-groups validity assesses a measure's ability to distinguish between groups of patients that differ according to some factor that is known or suspected to influence scores. The factors used in this study were patient perceived health status, patient perceived severity of RA, and how patients assessed their day. Mann-Whitney U tests (for 2 groups) and Kruskal-Wallis tests (for 3 or more groups) were employed to test for differences between these groups.

RESULTS

French-Canadian bilingual panel. The panel modified the instructions slightly to enhance clarity. For example, the word "statements" was translated as "sentences." Relatively few problems with the RAQoL items were encountered. Where a word-for-word translation was inappropriate, a phrase of equivalent meaning could easily be identified. For example, the phrase "a good cry" was translated as "a good crisis of tears" and "always on my mind" became "is always present in my soul." The greatest difficulty arose from the statement "I sometimes have problems using the toilet." The direct translation would imply problems with elimination or

mechanical problems with the toilet itself. A possible solution was to translate the item as, "I sometimes have difficulty making use of the toilet." Two alternative translations were produced for consideration by the lay panel: "I sometimes have difficulty using the toilet" and "I sometimes have difficulty making use of the toilet." Agreement could not be reached for a further 3 items and, again, 2 alternative translations were produced for each.

French-Canadian lay panel. Only minor changes were made in the instructions for the measure. For example "qu'une seule réponse" ("only one response") was changed to "une seule réponse" ("one response"). It was thought that the omission of "qu" would preserve the intended meaning while being more in line with familiar language usage. The panel members were easily able to choose between the alternative translations that had been produced for 4 of the items. The wording for all other items was considered acceptable.

English-Canadian lay panel. Only minor changes were made to the UK instructions to make them suitable in Canadian English. For example, "tick" became "check" and "at this moment" became "at this time." Similarly, certain words in the items themselves had to be changed due to the

differences in colloquial expressions used in the UK and Canada. For example, "shops" became "stores"; "condition" became "arthritis"; "jobs about the house" became "household chores," and "to keep stopping" became "frequently have to stop." The greatest difficulty was encountered with the statement "I feel tired whatever I do." This phrase had a different meaning to the panel members than that intended by the original item. In the UK, this item reflects tiredness regardless of what the individual is doing, whereas in Canada, it reflects tiredness regardless of what the individual is doing to prevent the tiredness. Thus, the item was modified to "Any amount of activity I do makes me feel tired." Modifications were agreed for every item with the exception of "I sleep badly at night." Panel members were unable to agree on whether "badly" should be replaced with "poorly." It was decided to retain both for consideration during the field-testing stage.

Field-testing. The demographic characteristics of the French-Canadian and English-Canadian samples are shown in Table 1.

French-Canadian field-testing. The RAQoL took between 4 and 15 minutes (mean 7.3 min) to complete. A majority of

Table 1. Demographic information for field testing and postal survey samples and perceived health status for postal survey.

	Field Test		Postal Survey	
	French	English	French	English
Sex				
Males, n (%)	3 (20.0)	2 (13.3)	12 (23.1)	15 (27.8)
Age				
Range, yrs	28–76	35–83	33–87	25–84
Mean (SD), yrs	52.4 (12.5)	61.3 (13.9)	60.5 (13.6)	64.0 (13.9)
Median (IQR), yrs	53.4 (46–60)	66.2 (51–72)	61.0 (50–71)	67.0 (55–75)
Marital status				
Married or living as, n (%)	6 (40.0)	9 (60.0)	25 (48.1)	29 (53.7)
Employment status				
Employed full-time, n (%)	4 (26.7)	2 (13.3)	9 (17.3)	8 (15.7)
Employed part-time, n (%)	3 (20.0)	2 (13.3)	2 (3.8)	1 (2.0)
Duration of illness				
Range, yrs	2.5–30	4–30	2–37	2–63
Mean (SD), yrs	13.5 (9.1)	15.8 (8.6)	12.9 (9.4)	16.9 (12.1)
Median (IQR), yrs	10.0 (8–20)	13.0 (10–23)	11.0 (6–16)	15.0 (10–22)
Perceived general health				
Excellent/very good, n (%)			2 (3.9)	2 (3.8)
Good, n (%)			15 (29.4)	22 (41.5)
Fair, n (%)			25 (49.0)	25 (47.2)
Poor, n (%)			9 (17.6)	4 (7.5)
Severity of condition				
Mild, n (%)			6 (11.8)	9 (17.0)
Moderate, n (%)			23 (45.1)	27 (50.9)
Severe, n (%)			20 (39.2)	13 (24.5)
Very severe, n (%)			2 (3.9)	4 (7.5)
Rating good day/bad day				
Very good, n (%)			2 (3.9)	4 (8.0)
Good, n (%)			31 (60.8)	37 (74.0)
Bad, n (%)			16 (31.4)	8 (16.0)
Very bad, n (%)			2 (3.9)	1 (2.0)

the respondents stated that they found the measure easy to understand and that the questions were clear and straightforward. Most reported that they could easily relate to the content of the questions. A few patients found the dichotomous yes or no response format to be restrictive: "Cannot always be seen as black or white but should be seen in terms of gray." To overcome this problem, it was decided to make the instruction on the questionnaire asking respondents to answer the questions in terms of how they feel at this time more prominent.

The items in the measure were well received with the exception of "I avoid all physical contact." Four of the 15 interviewees had difficulty with the meaning of this statement. One participant initially thought of sexual contact but then also considered sudden or surprise contact. For another participant, physical contact implied rough contact, for example, pushing or punching. Of the 4 respondents who reported difficulty with this question, only one changed the answer given when informed of the intended meaning of the item. Therefore, it was not considered necessary to change the wording of any of the items.

English-Canadian field testing. The time required to complete the RAQoL ranged from 4 to 15 minutes (mean 7 min). Again, the measure was well received, with interviewees judging it to be relevant and easy to understand and complete. All participants preferred "I sleep poorly at night" to "I sleep badly at night." As with the French-Canadian respondents, some interviewees reported that the yes/no response format was too restrictive. Once again, it was considered that emphasizing the "at this time" instruction would help overcome this criticism.

Postal surveys. Of the 92 French-Canadian patients surveyed, 64 (69.6%) responded. Of these, 12 were excluded as 10 of the questionnaires had incomplete information (one page was missing due to a photocopy error) and 2 respondents did not have RA. This resulted in an overall participation rate of 56.5%. Of the 87 English-Canadian patients surveyed, 61 (70.1%) responded. Of these, 7 were excluded: 3 patients had language problems, one did not have RA, one questionnaire had been completed by a relative, one was a mailing error, and one was not received in time for the analysis. This resulted in an overall participation rate of 62.1%. Therefore, 52 French-Canadian questionnaires and 54 English-Canadian questionnaires were completed for the 2 time points. The demographic characteristics of the 2 samples were similar, but there was a tendency for the French-Canadian sample to report worse perceptions of their health status in terms of general health, severity of RA, and rating of good day/bad day (Table 1).

Table 2 shows the summary descriptive statistics for the RAQoL and HAQ. Scores on the RAQoL can range from 0 to 30, with a high score indicating worse QoL. The HAQ is scored from 0 to 3, with a high score indicating worse disability. The percentage of cases obtaining the minimum

or maximum possible scores on the RAQoL was low, indicating minimal floor and ceiling effects. The end effects associated with the HAQ were higher than for the RAQoL, but still relatively low.

The number of respondents who missed at least one item on the RAQoL was higher than might be expected (Table 3). However, similar or higher levels were observed for the HAQ. The overall level of missing data was low for both versions of the RAQoL and was, again, similar or lower than that found with the HAQ. As the study served as a validation exercise, all cases with missing data were excluded from subsequent analyses.

Test-retest reliability. Table 4 shows the results of the assessment of the test-retest reliability of the 2 Canadian versions of the RAQoL. It can be seen that the reliability correlation coefficient exceeds the minimum requirement of 0.85 in both cases, indicating low levels of random measurement error. The stability of scores is further confirmed by the similarity of median scores at each time point.

Internal consistency. The Cronbach's alpha coefficients for the French-Canadian RAQoL were 0.92 at both time points. For the English-Canadian version, they were 0.93 on each administration. These results indicate that the constituent items of the RAQoL are adequately interrelated.

Construct validity. The correlation coefficient between the RAQoL and the HAQ at Time 1 was 0.77 for the French-speaking sample and 0.68 for the English-speaking sample. These moderately high correlations indicate the level of association that would be expected between a measure of QoL and one of disability in a disease such as RA, which has significant physical manifestations. Tables 5A and 5B show the results of the known-groups assessment for each version of the RAQoL at Time 1. It can be seen that both versions were able to distinguish groups defined in terms of perceived health status and perceived severity of RA, with the worst RAQoL scores statistically significantly associated with the most severe assessment group. Patients who rated today as bad or very bad obtained statistically significantly worse scores only on the French-Canadian RAQoL. The English speaking "bad" or "very bad" rating group had a higher (worse) score than the "very good" or "good" group, but this failed to reach statistical significance. It is possible that this may be due to the relatively small number of English-speaking patients who rated today as bad or very bad.

DISCUSSION

This study provides evidence that the French-Canadian and English-Canadian adapted versions of the RAQoL compare well to the original UK and Dutch versions. The new instruments represent valid and reliable tools for measuring quality of life in Canadian patients with RA.

The methodology of this study strictly adhered to the recommendations for the adaptations of the RAQoL for use in other languages and cultures. The use of this method-

Table 2. RAQoL, and HAQ descriptive scores.

	French		English	
	Time 1	Time 2	Time 1	Time 2
RAQoL				
n	39	40	41	38
Median (IQR)	16.0 (8.0–20.0)	17.0 (10.5–22.0)	13.0 (5.5–20.0)	10.5 (6.8–19.3)
Range	1–30	1–30	0–29	0–29
% at floor (0)	0	0	4.9	5.3
% at ceiling (30)	2.6	2.5	0	0
HAQ				
n	44	36	47	44
Median (IQR)	1.3 (0.6–1.8)	1.4 (0.7–1.8)	1.0 (0.5–1.6)	1 (0.6–1.5)
Range	0–2.5	0–2.6	0–2.8	0–2.8
% at floor (0)	9.1	5.6	4.3	9.1
% at ceiling (3)	0	0	0	0

Table 3. Missing data in the RAQoL and HAQ.

	French, n = 52		English, n = 54	
	Time 1	Time 2	Time 1	Time 2
RAQoL				
% cases with any missing responses	25.0	20.0	24.1	30.9
Overall % missing responses	2.3	3.0	2.0	3.6
HAQ				
% cases with any missing responses	32.7	28.0	27.8	21.8
Overall % missing responses	5.3	3.0	4.1	4.5

Table 4. Test-retest reliability of the RAQoL.

RAQoL Version	n	Median Scores (IQR)		Test-retest Correlation
		Time 1	Time 2	
French-Canadian	34	16.0 (7.8–20.3)	17.0 (10.0–22.0)	0.87
English-Canadian	32	11.5 (5.3–18.5)	9.5 (5.3–19.8)	0.95

Table 5A. Known-groups validity of the French-Canadian RAQoL.

	Assessment Group	n	Median (IQR)	p
Perceived health status	Excellent/good	15	7.0 (3–16)	< 0.001
	Fair/poor	24	19.0 (14–23)	
Severity of RA	Mild/moderate	22	11.5 (4–17)	< 0.001
	Severe/very severe	17	20.0 (15–27)	
Good day/bad day	Very good/good	26	14.5 (6–17)	< 0.05
	Bad/very bad	13	20.0 (15–25)	

Table 5B. Known-groups validity of the English-Canadian RAQoL.

	Assessment Group	n	Median (IQR)	p
Perceived health status	Excellent/good	19	8.0 (4–15)	< 0.005
	Fair/poor	22	19.0 (12–23)	
Severity of RA	Mild/moderate	28	8.5 (5–15)	< 0.001
	Severe/very severe	12	20.0 (15–25)	
Good day/bad day	Very good/good	32	10.5 (5–20)	NS
	Bad/very bad	7	19.0 (19–23)	

ology provides assurance that the intended meaning of the content in the original version was maintained throughout.

The RAQoL assesses needs-based QoL and can be seen as a valuable complement to the HAQ, an established measure of disability in RA. The use of both measures in clinical trials, together with appropriate assessments of impairments, will ensure that all aspects of outcome in RA are assessed. The study provides initial evidence that the RAQoL will be sensitive to real changes in quality of life over time. However, it remains necessary to formally evaluate the property of responsiveness⁸.

ACKNOWLEDGMENT

We are grateful to Jean François Baladi of Novartis Pharma Inc. for his support. We thank Dr. Michel Zummer, Dr. Jean-Pierre Mathieu, and Dr. Suzanne Mercille for providing us with French-Canadian patients, and Dr. Craig Watts, Dr. Hyman Tannenbaum, Dr. Michael Starr, and Dr. Martin Cohen for providing us with English-Canadian patients for the postal survey.

REFERENCES

1. Whalley D, McKenna SP, de Jong Z, van der Heijde D. Quality of life in rheumatoid arthritis. *J Rheumatol* 1997;36:884-8.
2. de Jong Z, van der Heijde D, McKenna SP, Whalley D. The reliability and construct validity of the RAQoL: A rheumatoid arthritis-specific quality of life instrument. *Br J Rheumatol* 1997;36:878-83.
3. Fries JF, Spitz P, Kraines RG, Holman H. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1988;31:471-9.
4. Hunt SM, Alonso J, Bucquet D, et al. Cross-cultural adaptation of health measures. *Health Policy* 1991;19:33-44.
5. Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: statistical and judgmental approaches. *J Rheumatol* 1982;9:753-7.
6. Weiner EA, Stewart BJ. *Assessing individuals*. Boston: Little Brown; 1984.
7. Streiner D, Norman G. *Health measurement scales*. Oxford: Oxford University Press; 1995.
8. Fortin PR, Stucki G, Katz JN. Measuring relevant change: An emerging challenge in rheumatologic clinical trials [editorial]. *Arthritis Rheum* 1995;38:1027-30.