

*Original article*

## JOA Back Pain Evaluation Questionnaire: initial report

The Clinical Outcomes Committee of the Japanese Orthopaedic Association, The Subcommittee on Evaluation of Back Pain and Cervical Myelopathy

The Subcommittee on Low Back Pain and Cervical Myelopathy Evaluation of the Clinical Outcome Committee of the Japanese Orthopaedic Association

MITSURU FUKU<sup>1</sup>, KAZUHIRO CHIBA<sup>2</sup>, MAMORU KAWAKAMI<sup>3</sup>, SHINICHI KIKUCHI<sup>4</sup>, SHINICHI KONNO<sup>4</sup>, MASABUMI MIYAMOTO<sup>5</sup>, ATSUSHI SEICHI<sup>6</sup>, TADASHI SHIMAMURA<sup>7</sup>, OSAMU SHIRADO<sup>8</sup>, TOSHIHIKO TAGUCHI<sup>9</sup>, KAZUHISA TAKAHASHI<sup>10</sup>, KATSUSHI TAKESHITA<sup>6</sup>, TOSHIKAZU TANI<sup>11</sup>, YOSHIKI TOYAMA<sup>2</sup>, EIJI WADA<sup>12</sup>, KAZUO YONENOBU<sup>13</sup>, TAKASHI TANAKA<sup>14</sup>, and YOSHIO HIROTA<sup>15</sup>

<sup>1</sup>Laboratory of Statistics, Osaka City University Faculty of Medicine, Osaka, Japan

<sup>2</sup>Department of Orthopaedic Surgery, Keio University, Tokyo, Japan

<sup>3</sup>Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan

<sup>4</sup>Department of Orthopaedic Surgery, School of Medicine, Fukushima Medical University, 1 Hikariga-oka, Fukushima 960-1295, Japan

<sup>5</sup>Department of Orthopedic Surgery, Nippon Medical School, Tokyo, Japan

<sup>6</sup>Department of Orthopaedic Surgery, the University of Tokyo, Tokyo, Japan

<sup>7</sup>Department of Orthopaedic Surgery, Iwate Medical University School of Medicine, Morioka, Japan

<sup>8</sup>Department of Orthopaedic Surgery, Saitama Medical School, Iruma-gun, Japan

<sup>9</sup>Department of Orthopedic Surgery, Yamaguchi University School of Medicine, Yamaguchi, Japan

<sup>10</sup>Department of Orthopedic Surgery, Graduate School of Medicine, Chiba University, Chiba, Japan

<sup>11</sup>Department of Orthopaedics, Kochi Medical School, Kochi, Japan

<sup>12</sup>Department of Orthopaedic Surgery, Hoshigaoka Koseinenkin Hospital, Osaka, Japan

<sup>13</sup>National Hospital Organization Osaka-Minami Medical Center, Kawachinagano, Japan

<sup>14</sup>Department of Internal Medicine, Houai Hospital, Osaka, Japan

<sup>15</sup>Department of Public Health, Osaka City University Faculty of Medicine, Osaka, Japan

### Abstract

**Background.** There is no widely accepted objective evaluation for lumbar spine disorders. New outcome measures should be patient-oriented and should measure symptoms and self-reported functional status in multiple dimensions. The aim of this study was to identify items to be included in the disease-specific quality of life (QOL) questionnaire for the assessments of patients with lumbar spine disorders.

**Methods.** The draft of the QOL questionnaire that consisted of a total of 60 items, including 24 items derived from the Japanese version of the Roland Morris Disability Questionnaire (RDQ) and 36 items derived from the Japanese version of Short Form 36 (SF-36), were administered to patients and controls. After obtaining written informed consent, the following data were collected from the patient group ( $n = 328$ ) and the control group ( $n = 213$ ): (1) background characteristics, including age, diagnosis, Japanese Orthopaedic Association (JOA) score, and finger to floor distance; (2) responses to the questionnaire; (3) the identification rate by discrimination analysis to select the candidates for adoption and by adopting explanatory variables. The items to be excluded were determined by examining the explanatory variables, which were

selected after the discrimination analysis, by setting the candidate to-be-excluded items as an objective variable.

**Results.** Based on the distribution of the responses, two items, RDQ-15 and RDQ-19, were excluded. From the results of the correlation coefficient calculation for each question in the patient group, 33 items were excluded and 27 candidate items were adopted. Based on the adoption explanatory variable used in the discrimination analysis, 25 of the 27 candidate items for adoption were accepted.

**Conclusions.** This study identified the 25 specific questionnaire items that should be included in the questionnaire to evaluate QOL of patients with various lumbar spine disorders.

### Introduction

The assessments of lumbar spine disorders have been based on biological, physiological, and anatomical outcomes, such as measurements of the range of spinal motion, laboratory tests, and imaging studies.<sup>1</sup> However, these indicators have little meaning for the patient and the society. On the other hand, alleviation of symptoms, such as pain intensity, and an improved quality of life

Offprint requests to: S. Konno

Received: December 18, 2006 / Accepted: June 19, 2007

(QOL) have more significance for the patients and the society. It has been reported that patient self-rated measures of symptom intensity and QOL are as reproducible as many physiological measurements and are acceptable with respect to objectivity and stability.<sup>2</sup> Thus, patient-based outcomes involving patient self-assessment of symptom intensity and QOL should be used in clinical research.

Conventionally, surgery is evaluated based on a simple four-grade scale: excellent, good, fair, and poor. This approach has limitations due to its subjectivity and the lack of clear definitions for each grade. Therefore, the evaluation of treatment results depends on individual researchers and is not fully comparable. Furthermore, the four-grade scale is not sufficient to measure pain intensity, activities of daily living, or the ability to work. For example, a patient might not be able to return to work despite a decrease in pain, or there may be no alleviation of pain intensity despite an improvement in activities of daily living. Given such circumstances, an improvement in one dimension does not necessarily mean an improvement in other dimensions; thus, the evaluation of medical treatments must be multidimensional and include patient-based outcomes. Given these perspectives, the Assessment Standards Committee prepared this report dealing with the new standards for evaluating the results of treatments for lumbar spine disorders.

## Materials and methods

### *Selection of lumbar spine disorders evaluation items*

The aim of this study was to establish a multidimensional method for evaluating treatment results for lumbar spine disorders that was centered on patient-based outcomes and that could be used internationally. Pain intensity can be measured using a visual analogue scale (VAS) and the NASS questionnaire.<sup>3,4</sup> The Roland-Morris disability questionnaire (RDQ) and the Oswestry disability questionnaire are low back pain-specific QOL questionnaires.<sup>5,6</sup> With respect to the RDQ, the Japanese version of the RDQ has been developed that conforms to the psychometric standards in the areas of reliability, validity, and responsiveness.<sup>7</sup> Both alleviation of patients' symptoms and its impact on their activities of daily living can be measured using the RDQ. Widely used international measures for general well-being include the SF36, SF12, and the Euro QOL; Japanese versions of the SF36 and the Euro QOL have been developed.<sup>8-11</sup> Thus, it would be desirable to use the VAS for measuring the intensity of low back pain, the RDQ for measuring low back pain-specific QOL, and the SF-36 for assessing general well-being. However,

the evaluation of all of these items in daily practice is impractical owing to the large number of items. The approximate time to complete the RDQ is 5 min, and it takes 10 min to complete the SF36.<sup>1</sup> To reduce the number of items necessary to evaluate the efficacy of treatments for lumbar spine disorders, the usefulness of various evaluation criteria to differentiate patients with lumbar spine disorders from normal subjects was studied.

### *Examination of the evaluation rating score (true value) in the lumbar spine disorders group*

Eight institutions (including affiliated institutions) were asked to recruit at least 40 subjects during the period from February to May 2002. The questionnaire consisted of a total of 60 items: 24 items derived from the Japanese version of the RDQ and 36 items derived from the Japanese version of the SF-36. Lumbar disc herniation and lumbar canal stenosis were the two main targets. Subjects who had other orthopedic disorders and those with impaired ability to understand the questions, such as patients with dementia, were excluded. Normal subjects were defined as adults with no orthopedic disorders. Adults living independently and not requiring nursing care but who were undergoing alternative treatments (e.g., acupuncture, moxibustion, massage, and chiropractic treatments) were included in the control group. Health care professionals were excluded.

Prior to conducting the investigations, subjects in the patient group and the control group gave their written informed consent.

### *Background characteristics of the patient group*

The distribution of subjects' background characteristics, such as age, diagnosis, Japanese Orthopaedic Association (JOA) score, and finger to floor distance, was analyzed to verify that the group represents the general population of patients with spine disorders.

### *Examination of removable candidate QOL items*

A QOL item could be removed if it satisfied any of the following criteria: (1) items to which most subjects gave the same answer; (2) items the answers for which were highly correlated with the answers to other questions; (3) items that could be explained by several questions; (4) items whose score distributions did not show any statistically significant differences between the patient and control groups.

To test for the above conditions, the distribution of responses for the RDQ and SF36 were compared between the two groups. The correlation coefficient for each question in the patient group was analyzed using the Spearman correlation coefficient.

**Table 1.** Demographics of patients and controls

	Patients			Controls		
	Male	Female	Total	Male	Female	Total
	187	159	346	96	120	216
Age group (years)						
10-19	1	2	3			
20-29	19	16	35	15	33	48
30-39	23	13	36	18	22	40
40-49	18	19	37	20	16	36
50-59	30	29	59	15	21	36
60-69	39	40	79	12	15	27
70-79	47	32	79	12	12	24
80-89	9	8	17	4	1	5
≥90	1		1			
Diagnosis						
Lumbar disc herniation (LDH)			160			
Lumbar spinal canal stenosis (LCS)			183			
LDH + LCS			3			
Other orthopedic disorders						
Present			18			3
Absent			328			213

#### Examination of the identification rate by discrimination analysis of candidate items

After using the above-described criteria to identify the candidate items to be included in the final questionnaire, discrimination analysis was done to eliminate further the number of items. By setting one of the candidate items for adoption as the objective variable, the rest of the items were examined as explanatory variables; the discrimination rate was then analyzed, and items with a minimum discrimination rate  $\geq 70\%$  were considered to be items that could be excluded. The final items that were excluded were determined by examining the explanatory variables, which were selected after discrimination analysis, setting the candidate to-be-excluded items as the objective variable.

## Results

#### Background characteristics of the patient group

Table 1 shows the age, sex, and diagnosis of 328 subjects in the patient group and 213 subjects in the control group. There was significant difference in sex and age distribution between the two groups ( $P = 0.03$ , Fisher's exact test). In the patient group, the straight leg raising (SLR) test was positive in approximately 40%, sensory disturbance was present in 60%, muscle weakness was seen in 40%, and bladder dysfunction was impaired in approximately 10% of the subjects (Table 2). The distribution of the finger to floor distance revealed that the

**Table 2.** Clinical findings

Parameter	Patients	Control
	328	213
SLR		
Normal	201	211
30°-70°	110	1
<30°	17	1
Sensory		
Normal	124	210
Mild disturbance	150	2
Obvious disturbance	54	1
Motor		
Normal	191	212
Mild muscle weakness	107	0
Obvious muscle weakness	30	1
Bladder function		
Normal	293	207
Mild dysuria	33	6
Severe dysuria	2	0
Finger-to-floor distance		
- -15	1	6
-14 - -5	3	10
-4 - 4	95	108
5 - 14	68	54
15 - 24	54	30
25 - 34	51	5
35 - 44	17	0
45 - 54	24	0
55 - 64	8	0
65 - 74	1	0
75 - 84	2	0
Measurement not possible	4	0

SLR, straight-leg raising

mobility of the lumbar spine in the patient group was significantly restricted compared to that of the control group. Although we cannot make any conclusions, given the above results we considered that the patient group represented the general population of the patients with lumbar spine disorders.

### RDQ

The nonresponse rate for the RDQ was less than 5% for all questions; no questions were difficult to answer. As expected, more than 95% of the normal subjects answered "no" to all questions. In the patient group,

**Table 3.** Results of the RDQ (Roland-Morris Disability Questionnaire)

Question item	Patients (n = 328)			Controls		
	Yes	No	No response	Yes	No	No response
RDQ-1	174 53.0%	148 45.1%	6 1.8%	1 0.5%	212 99.5%	
RDQ-2	232 70.7%	92 28.0%	4 1.2%	20 9.4%	193 90.6%	
RDQ-3	253 77.1%	69 21.0%	6 1.8%	7 3.3%	204 95.8%	2 0.9%
RDQ-4	110 33.5%	210 64.0%	8 2.4%	2 0.9%	211 99.1%	
RDQ-5	183 55.8%	135 41.2%	10 3.0%	4 1.9%	209 98.1%	
RDQ-6	215 65.5%	109 33.2%	4 1.2%	3 1.4%	208 97.7%	2 0.9%
RDQ-7	119 36.3%	204 62.2%	5 1.5%	3 1.4%	210 98.6%	
RDQ-8	185 56.4%	139 42.4%	4 1.2%	4 1.9%	209 98.1%	
RDQ-9	152 46.3%	170 51.8%	6 1.8%	2 0.9%	211 99.1%	
RDQ-10	193 58.8%	129 39.3%	6 1.8%	4 1.9%	209 98.1%	
RDQ-11	166 50.6%	156 47.6%	6 1.8%	6 2.8%	207 97.2%	
RDQ-12	86 26.2%	237 72.3%	5 1.5%	1 0.5%	211 99.1%	1 0.5%
RDQ-13	128 39.0%	195 59.5%	5 1.5%	11 5.2%	202 94.8%	
RDQ-14	162 49.4%	159 48.5%	7 2.1%	2 0.9%	211 99.1%	
RDQ-15	53 16.2%	270 82.3%	5 1.5%		212 99.5%	1 0.5%
RDQ-16	166 50.6%	157 47.9%	5 1.5%	3 1.4%	209 98.1%	1 0.5%
RDQ-17	222 67.7%	100 30.5%	6 1.8%	6 2.8%	206 96.7%	1 0.5%
RDQ-18	114 34.8%	210 64.0%	4 1.2%	2 0.9%	210 98.6%	1 0.5%
RDQ-19	28 8.5%	296 90.2%	4 1.2%		213 100.0%	
RDQ-20	109 33.2%	211 64.3%	8 2.4%	1 0.5%	212 99.5%	
RDQ-21	263 80.2%	58 17.7%	7 2.1%	7 3.3%	206 96.7%	
RDQ-22	105 32.0%	219 66.8%	4 1.2%	4 1.9%	209 98.1%	
RDQ-23	257 78.4%	65 19.8%	6 1.8%	10 4.7%	203 95.3%	
RDQ-24	82 25.0%	242 73.8%	4 1.2%		213 100.0%	

**Table 4.** Exclusion and adoption of items (first level)

Excluded items	Reason	Items adopted
RDQ-3	RDQ-17, 23 Correlation	RDQ-1
RDQ-7	RDQ-12 Correlation	RDQ-2
RDQ-9	RDQ-16 Correlation	RDQ-4
RDQ-10	RDQ-1, 17 Correlation	RDQ-5
RDQ-15	Answers concentrated on [NO]	RDQ-6
RDQ-19	Answers concentrated on [NO]	RDQ-8
RDQ-21	RDQ-17, 23 Correlation	RDQ-11
RDQ-24	RDQ-4 Correlation	RDQ-12
QOL-2	QOL-1 Correlation	RDQ-13
QOL-3a	QOL-3g Correlation	RDQ-14
QOL-3b	QOL-3g Correlation	RDQ-16
QOL-3c	QOL-3e Correlation	RDQ-17
QOL-3d	QOL-3e, 3g Correlation	RDQ-18
QOL-3h	QOL-3e, 3g Correlation	RDQ-20
QOL-3i	QOL-3e, 3g Correlation	RDQ-22
QOL-3j	QOL-3e Correlation	RDQ-23
QOL-4a	QOL-8 Correlation	QOL-1
QOL-4c	QOL-8 Correlation	QOL-3e
QOL-4d	QOL-8 Correlation	QOL-3f
QOL-5a	QOL-8 Correlation	QOL-3g
QOL-5b	QOL-8, 9f Correlation	QOL-4b
QOL-5c	QOL-8, 9f Correlation	QOL-8
QOL-6	QOL-8 Correlation	QOL-9f
QOL-7	QOL-8 Correlation	QOL-9g
QOL-9a	QOL-9h Correlation	QOL-9h
QOL-9b	QOL-9f Correlation	QOL-11b
QOL-9c	QOL-9f Correlation	QOL-11c
QOL-9d	QOL-9f, 9h Correlation	
QOL-9e	QOL-9h Correlation	
QOL-9i	QOL-9g Correlation	
QOL-10	QOL-8 Correlation	
QOL-11a	QOL-11b Correlation	
QOL-11d	QOL-1, 11b Correlation	

more than 80% of respondents chose the same answers for items 15, 19, and 21; and approximately 80% chose the same answer for items 3 and 23. In particular, for items 15 and 19, more than 80% of the patient group chose the same answer (no) as the normal healthy subjects (Table 3). Therefore, based on these results, RDQ-15 and RDQ-19 were listed as candidates to be excluded.

#### SF-36

The nonresponse rate for the SF-36 was less than 5% for all questions, and none of the questions was difficult to answer. There was a statistically significant difference in the distribution of responses between the patient group and the control group ( $P < 0.05$ , by  $\chi^2$ ). Furthermore, there were no questions for which the answers were predominantly concentrated on one choice in the patient group.

#### Correlation coefficient for each question in the patient group

For the 24 RDQ items, there were mutual correlations between two groups of items: RDQ-1, 3, 10, 17, 21, and 23 (six items); and RDQ-4, 7, 9, 12, 16, and 24 (six items). For the SF-36 items, there were mutual correlations among four groups: QOL-1, 2, 11a, 11b, and 11d (5 items); QOL-3a–3j (10 items); QOL-4a–4d, 5a–5c, 6, 7, 8, and 10 (11 items); QOL-9a–9i (9 items). Thus, 33 items were excluded, and 27 remained as candidates for adoption. The reasons for exclusion are shown in Table 4.

#### Discrimination analysis

The discrimination rate of the answer for each item, based on the discrimination analysis, was determined for the 27 candidates for adoption. To arrive at the discrimination rate, one item was set as the objective variable, and the other items were set as explanatory variables; the item with a high minimum value for the

**Table 5.** Results of discrimination analysis

Question item	Minimal ratio of discrimination for each choice	Ratio of discrimination through all choices	$\kappa$
RDQ-1	75.2% <sup>a</sup>	77.3% <sup>a</sup>	0.54
RDQ-2	60.5%	81.5% <sup>a</sup>	0.53
RDQ-4	66.7%	81.0% <sup>a</sup>	0.57
RDQ-5	74.8% <sup>a</sup>	80.1% <sup>a</sup>	0.59
RDQ-6	62.1%	80.5% <sup>a</sup>	0.55
RDQ-8	69.2%	76.7% <sup>a</sup>	0.52
RDQ-11	69.0%	72.7% <sup>a</sup>	0.45
RDQ-12	69.6%	87.0% <sup>a</sup>	0.65
RDQ-13	55.0%	68.8%	0.34
RDQ-14	70.6% <sup>a</sup>	74.5% <sup>a</sup>	0.49
RDQ-16	76.6% <sup>a</sup>	78.3% <sup>a</sup>	0.57
RDQ-17	64.2%	83.4% <sup>a</sup>	0.59
RDQ-18	60.0%	77.2% <sup>a</sup>	0.48
RDQ-20	61.0%	77.0% <sup>a</sup>	0.47
RDQ-22	44.4%	72.6% <sup>a</sup>	0.33
RDQ-23	57.1%	84.3% <sup>a</sup>	0.50
QOL-1	30.0%	56.2%	0.31
QOL-3e	52.8%	68.8%	0.45
QOL-3f	46.3%	61.5%	0.38
QOL-3g	50.0%	69.7%	0.50
QOL-4b	48.1%	58.2%	0.46
QOL-8	46.0%	55.9%	0.41
QOL-9f	36.7%	56.6%	0.41
QOL-9g	28.6%	47.1%	0.28
QOL-9h	14.3%	48.5%	0.28
QOL-11b	21.1%	48.3%	0.27
QOL-11c	5.6%	58.7%	0.23

<sup>a</sup>Discrimination rate >70%

discrimination rate was excluded from adoption. The minimum value for the discrimination rate was > 70% in four items (RDQ-1, 5, 14, and 16) (Table 5). The discrimination rate calculated the ratio that the answers of patients group accorded with the estimated answers by classification rule. To compute the  $\kappa$  value, we made a contingency table using the answers of patients group and by the estimated answers.

#### *Adoption of the explanatory variables in discrimination analysis*

To verify whether it would be appropriate to exclude RDQ-1, 5, 14, and 16, the explanatory variable chosen for each objective variable in discrimination analysis was determined (Table 6). Consequently, it was found that if RDQ-1 and 5 were excluded RDQ-14 and 16 could not be excluded because RDQ-1 and 5 would be necessary. Given these results, 25 of the 27 candidate items for adoption were adopted; RDQ-1 and 5 were excluded.

#### **Discussion**

Several issues must be considered when creating a new evaluation method for medical treatments. First, the evaluation should be structured so the effect of medical intervention is accurately reflected. If medical treatment results are mainly determined by genetic or environmental factors, the quality of the treatment cannot be evaluated. Second, the evaluation of medical treatment results must contain a framework that accurately and reliably captures changes in the patient's health condition. Finally, to evaluate the medical treatment results accurately, the treatment evaluation period should be the same as the time period during which information is obtained about the patients' complications and social background that can affect the medical treatment outcomes.

Evaluation of medical treatment outcomes used to be a subject of concern for health care professionals only. Recently, however, the evaluation of medical treatment outcomes is becoming more of a concern to patients and governments who pay the medical costs. Evaluating medical treatment results is key to assessing cost effectiveness and to validating treatments themselves. Thus,

**Table 6.** Explanatory variable chosen for every objective variable on discrimination analysis

		Objective variable			
		RDQ-1	RDQ-5	RDQ-14	RDQ-16
Minimum value of discrimination rate for every choice		75.2%	74.8%	70.6%	76.6%
All choices by discrimination rate		77.3%	80.1%	74.5%	78.3%
$\kappa$		0.54	0.59	0.49	0.57
Variable	Explanation				
RDQ-1	I stay at home most of the time because of my back.	-		○	○
RDQ-2	I change position frequently to try to get my back comfortable.			○	○
RDQ-4	Because of my back, I am not doing any of the jobs that I usually do around the house.	○	○	○	
RDQ-5	Because of my back, I use a handrail to get upstairs.		-		○
RDQ-6	Because of my back, I lie down to rest more often.	○			○
RDQ-8	Because of my back, I try to get other people to do things for me.	○			○
RDQ-11	Because of my back, I try not to bend or kneel down.	○		○	○
RDQ-12	I found it difficult to get out of a chair because of my back.		○	○	○
RDQ-13	My back is painful almost all the time.		○	○	
RDQ-14	I find it difficult to turn over in bed because of my back.	○		-	○
RDQ-16	I have trouble putting on my socks (or stockings) because of the pain in my back.	○		○	-
RDQ-17	I only walk short distances because of my back.	○	○	○	
RDQ-18	I sleep less well because of my back.		○	○	
RDQ-20	I sit down for most of the day because of my back.	○	○		○
RDQ-22	Because of my back pain, I am more irritable and bad tempered with people than usual.	○			○
RDQ-23	Because of my back, I go upstairs more slowly than usual.	○	○		○
QOL-1	In general, would you say your health is:				
QOL-3e	Does your health now limit you in climbing one flight of stairs?		○		
QOL-3f	Does your health now limit you in bending, kneeling, or stooping?		○		○
QOL-3g	Does your health now limit you in walking more than a mile?		○		○
QOL-4b	Any problems as a result of your physical health: accomplished less than you would like?				
QOL-8	How much did pain interfere with your normal work?	○			
QOL-9f	Have you felt downhearted and blue?				
QOL-9g	Did you feel worn out?				
QOL-9h	Have you been a happy person?	○			○
QOL-11b	I am as healthy as anybody I know.	○		○	○
QOL-11c	I expect my health to get worse.	○			

criteria used for the evaluations should be objective and structured in such a way that the patients' perspective is respected. In this way, the results can be understood not only by health care professionals but also by patients and third parties. Evaluation of medical treatment based on the creation of standards can be used to document and improve the performance of the medical system and health care technology.

This study has several limitations. There was a significant difference in sex and age between the patients group and the normal group. Hence, there is a possibility that this difference affects the results of our study. For many research purposes, it may be optimal to include both disease-specific (RDQ) and generic functional status measures (SF-36). However, an instrument that includes both disease-specific and general func-

tional status measures has not been established. Although it may not be ideal to combine items from two different instruments, our final goal was to find the disease-specific daily functions, physical function, role function, pain, vitality, mental health, and health perception. However, only the Japanese version of the RDQ as the disease-specific and the Japanese version of the SF-36 as the generic functional status measure were available. Therefore, we combined items from the two instruments to find the disease-specific functional status that included many dimensions.

We were able to identify 25 specific questions that would elucidate the QOL of patients with various lumbar spine disorders. The next step is to assess the validity and responsiveness of the questionnaire that includes the selected 25 questions by measuring the

outcome of patients with lumbar spinal disorders. Also, we have to complete cross-cultural adaptation of the BPEQ so it can be used internationally.

## References

1. Deyo RA, Weinstein JN. Outcomes research for spinal disorders. In: Herkowitz HN, Garfin SR, Balderston RA, Eismont FJ, Bell GR, Wiesel SW, editors. *Rothman-Simeone, the spine*. 4th edn. Philadelphia: Saunders; 1999. p. 173–83.
2. Deyo RA. Measuring the functional status of patients with low back pain. *Arch Phys Med Rehabil* 1988;69:1044–53.
3. Bond MR, Pilowsky I. Subjective assessment of pain and its relationship to the administration of analgesics in patients with advanced cancer. *J Psychosom Res* 1966;10:203–8.
4. Deyo RA, Andersson G, Bombardier C, Cherkin DC, Keller RB, Lee CK, et al. Outcome measures for studying patients with low back pain. *Spine* 1994;19:2032S–6S.
5. Roland M, Morris R. A study of the natural history of back pain. Part 1. Development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983;8:141–4.
6. Fairbank JCT, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
7. Suzukamo Y, Fukuhara S, Kikuchi S, Konno S, Roland M, Iwamoto Y, et al. Validation of the Japanese version of the Roland-Morris Disability Questionnaire. *J Orthop Sci* 2003;8:543–8.
8. Fukuhara S. QOL assessment in outcomes research. *Nihon Youtsukaishi (Journal of Japanese Society of Lumbar Spine Disorders)* 2002;8:31–7 (in Japanese).
9. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form survey (SF36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
10. Ware JE Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.
11. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine* 1995;20:1899–909.

## *Original article*

# Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire: Part 3. Determination of reliability

MITSURU FUKUI, KAZUHIRO CHIBA, MAMORU KAWAKAMI, SHINICHI KIKUCHI, SHINICHI KONNO, MASABUMI MIYAMOTO, ATSUSHI SEICHI, TADASHI SHIMAMURA, OSAMU SHIRADO, TOSHIHIKO TAGUCHI, KAZUHISA TAKAHASHI, KATSUSHI TAKESHITA, TOSHIKAZU TANI, YOSHIAKI TOYAMA, KAZUO YONENOBU, EIJI WADA, TAKASHI TANAKA, and YOSHIO HIROTA

Department of Orthopedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

### Abstract

**Background.** The manner of measuring the outcome of cervical myelopathy must be patient-oriented and have sufficient reliability and validity. The current Japanese Orthopaedic Association (JOA) scoring system for cervical myelopathy is widely used but has not met this requirement. The first- and second-round surveys established 24 items for inclusion on a new questionnaire for cervical myelopathy. The purpose of this study (the third-round survey A) was to confirm the reproducibility of patient responses to the selected questions.

**Methods.** A total of 201 patients with cervical myelopathy and with no change of symptoms between the two interviews were included. Each patient was interviewed twice using the same questionnaire at an interval of 4 weeks. The reliability of the questionnaire was evaluated by determining the extension of the weighted kappa coefficients.

**Results.** The weighted kappa coefficient for each item was  $>0.4$ , confirming that the test–retest reliability was acceptable.

**Conclusions.** The newly developed JOA Cervical Myelopathy Evaluation Questionnaire was proven to have sufficient reliability.

statistics, and (2) more patient-oriented. For the first- and second-round surveys, the committee created a new questionnaire of 77 items including SF-36 (the Medical Outcome Study Short-Form 36-Item Health Survey)<sup>3</sup> to assess health-related quality of life (HRQOL) and 41 items to assess functioning of the cervical spinal cord and cervical spine. In these surveys, 53 items were excluded.<sup>1,2</sup> The purpose of this study (the third-round survey A) was to confirm the reproducibility of the selected 24 questions written in Japanese (see Appendix). The questionnaire was self-administered, so interobserver reliability did not need to be confirmed. To examine the reliability of the questionnaire, only test–retest reliability was evaluated.

### Materials and methods

Inclusion criteria were as follows: (1) Patient had to have cervical myelopathy secondary to cervical disc herniation, cervical spondylosis, or ossification of the posterior longitudinal ligament (OPLL); and (2) there was no change of symptoms between the two interviews, which was judged by attending surgeons. Exclusion criteria included (1) patient had other musculoskeletal disease requiring medical treatment; (2) patient had a mental disease and so could potentially furnish inappropriate answers; (3) patient was postoperative; 4) patient had participated in the former surveys.

We asked 460 Japanese board-certified spine surgeons to participate in the survey between January 2004 and June 2004. Patients were recruited from the outpatient clinic of each surgeon during the planned inquiry period. Self-assessment was used in this study. Each patient was interviewed using the same questionnaire twice at an interval of 4 weeks ( $\pm 3$  days). The attending surgeon filled out the patient's information

### Introduction

As described in Part 1<sup>1</sup> and Part 2,<sup>2</sup> the Japanese Orthopaedic Association (JOA) decided to revise the assessment tools for cervical myelopathy and develop a new JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ). The point of the revision was to make it: (1) more scientific from the standpoint of medical

Offprint requests to: A. Seichi

Received: October 10, 2006 / Accepted: March 2, 2007

The authors are members of the Subcommittee on Low Back Pain and Cervical Myelopathy, Evaluation of the Clinical Outcome Committee of the Japanese Orthopaedic Association.

regarding the diagnosis and the presence or absence of concomitant diseases and judged the severity using a three-step rating scale (mild, moderate, severe). This study was approved by the Ethics Committee of the Japanese Society for Spine Surgery and Related Research, and informed consent was obtained from each patient.

The reliability of the questionnaire was evaluated by determining the extension of the weighted kappa coefficients. According to the most widely accepted criteria proposed by Landis and Koch, a kappa coefficient of  $\geq 0.4$  was judged reliable.<sup>4</sup> The 95% confidence intervals were calculated for all reliability coefficients.

All statistics were calculated using SPSS software (version 12; SPSS, Chicago, IL, USA).

**Results**

Of 304 patients who participated in this survey, 103 were excluded owing to a change of symptoms between the two interviews or an inappropriate patient administration interval, leaving 201 available for the study (74 women, 127 men). The mean  $\pm$  SD patient age was 63.8  $\pm$  12.8 years (Table 1). The diagnosis was cervical disc herniation in 29, cervical spondylotic myelopathy in 113, and OPLL/other in 59. Table 2 lists the severity of patient myelopathy evaluated by the current JOA scoring system and shows that the characteristics of the recruited patients were not specific. The distribution of answers for each item during the second interview is

**Table 1.** Baseline participant demographics (*n* = 201)

Parameter	Male	Female	Total
No. of patients	127	74	201
Age (years)			
30s	7	4	11
40s	12	8	20
50s	26	13	39
60s	41	13	54
70s	32	27	59
80s	9	9	18
Average	63	65.2	63.8
SD	12.4	13.5	12.8
Minimum	30	34	30
Maximum	89	84	89
Median	64	69	66
Diagnosis			
Cervical disc herniation	17	12	29
Cervical spondylosis	68	45	113
OPLL	39	16	55
Other	3	1	4

OPLL, ossification of the posterior longitudinal ligament

shown in Table 3, and neither ceiling nor floor effects were observed. The results for test-retest reliability are shown in Table 4. The weighted kappa coefficient for each item was  $>0.4$ , showing that the test-retest reliability of JOACMEQ was acceptable. Among the three groups with severity judged as mild, moderate, or severe, there was some difference in the weighted kappa coefficient for some items, but the difference was statistically not significant. The difference among the three diseases (cervical disc herniation, cervical spondylosis, OPLL) was also statistically not significant (data not shown).

**Table 2.** Distribution of the severity evaluated by the current JOA scoring system

Score	No.
<b>Motor function</b>	
Upper extremity	
0	1
1	10
2	37
3	90
4	63
Lower extremity	
0	1
0.5	0
1	21
1.5	22
2	40
2.5	20
3	40
4	57
<b>Sensory function</b>	
Upper extremity	
0	2
0.5	22
1	88
1.5	61
2	28
Trunk	
0	1
0.5	1
1	14
1.5	26
2	159
Lower extremity	
0	1
0.5	15
1	55
1.5	44
2	86
<b>Bladder function<sup>a</sup></b>	
0	4
1	17
2	52
3	127

JOA, Japanese Orthopaedic Association

<sup>a</sup>Data defect: *n* = 1

**Table 3.** Distribution of answers for each item in the second interview

Item	1	2	3	4	5	No answer
C01	108 53.70%	80 39.80%	13 6.50%			
C02	155 77.10%	41 20.40%	5 2.50%			
C03	124 61.70%	53 26.40%	17 8.50%	6 3.00%		1 0.50%
C04	105 52.20%	61 30.30%	23 11.40%	10 5.00%	2 1.00%	
C05	86 42.80%	74 36.80%	37 18.40%			4 2.00%
C06	132 65.70%	43 21.40%	15 7.50%	10 5.00%		1 0.50%
C07	65 32.30%	103 51.20%	33 16.40%			
C08	114 56.70%	70 34.80%	16 8.00%			1 0.50%
C09	129 64.20%	57 28.40%	14 7.00%			1 0.50%
C10	111 55.20%	60 29.90%	27 13.40%			3 1.50%
C11	121 60.20%	68 33.80%	12 6.00%			
C12	138 68.70%	52 25.90%	11 5.50%			
C13	68 33.80%	88 43.80%	44 21.90%			1 0.50%
QOL01	6 3.00%	16 8.00%	77 38.30%	90 44.80%	11 5.50%	1 0.50%
QOL02	39 19.40%	91 45.30%	69 34.30%			2 1.00%
QOL03	36 17.90%	97 48.30%	66 32.80%			2 1.00%
QOL04	33 16.40%	74 36.80%	93 46.30%			1 0.50%
QOL05	16 8.00%	29 14.40%	91 45.30%	44 21.90%	20 10.00%	1 0.50%
QOL06	15 7.50%	49 24.40%	67 33.30%	47 23.40%	22 10.90%	1 0.50%
QOL07	11 5.50%	26 12.90%	95 47.30%	43 21.40%	25 12.40%	1 0.50%
QOL08	12 6.00%	35 17.40%	92 45.80%	45 22.40%	15 7.50%	2 1.00%
QOL09	17 8.50%	46 22.90%	102 50.70%	30 14.90%	5 2.50%	1 0.50%
QOL10	8 4.00%	56 27.90%	65 32.30%	55 27.40%	17 8.50%	
QOL11	18 9.00%	60 29.90%	74 36.80%	38 18.90%	11 5.50%	

## Discussion

Measurements of the outcome of cervical myelopathy must have sufficient reliability and validity and should be proven by means of statistical evaluation. The current JOA scoring system for cervical myelopathy has been widely used but has not been shown to meet such a requirement.<sup>5,6</sup> A manner of measuring patient-based outcome has also been regarded as essen-

tial for evaluating a patient's health status. We sought to develop a new questionnaire using a psychometric method. The newly proposed document, which is self-administered and disease-specific, includes function of the cervical spinal cord and cervical spine as well as health-related quality of life. The first- and second-round surveys selected 24 of 77 items after repeated discussions among the subcommittee members.<sup>1,2</sup>

**Table 4.** Reproducibility of each item

Item	Item no.	$\kappa$	95% CI
C01	3	0.76	0.69–0.82
C02	3	0.59	0.50–0.67
C03	4	0.71	0.64–0.77
C04	5	0.74	0.69–0.79
C05	3	0.75	0.70–0.80
C06	5	0.69	0.63–0.76
C07	3	0.80	0.75–0.85
C08	3	0.65	0.58–0.71
C09	3	0.68	0.62–0.75
C10	3	0.63	0.57–0.69
C11	3	0.72	0.65–0.77
C12	3	0.55	0.47–0.62
C13	3	0.71	0.66–0.76
QOL01	3	0.57	0.51–0.63
QOL02	3	0.68	0.62–0.74
QOL03	3	0.64	0.58–0.70
QOL04	3	0.73	0.67–0.78
QOL05	5	0.58	0.52–0.63
QOL06	5	0.55	0.49–0.60
QOL07	5	0.62	0.56–0.67
QOL08	5	0.56	0.50–0.61
QOL09	5	0.56	0.50–0.61
QOL10	5	0.62	0.57–0.67
QOL11	5	0.49	0.42–0.54

In this third survey, we succeeded in confirming the reliability of the 24 items. The next steps are to check the validity of the questionnaire, perform a factor analysis to determine the underlying cluster of the question-

naire items, complete the measurement scale, and confirm the responsiveness of the questionnaire.

### Conclusion

The newly developed JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ) was proven to be sufficiently reliable.

### References

1. Fukui M, Chiba K, Kawakami M, Kikuchi S, Konno S, Miyamoto M, et al. An outcome measure for patients with cervical myelopathy: Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ). Part 1. *J Orthop Sci* 2007;12:227–40.
2. Fukui M, Chiba K, Kawakami M, Kikuchi S, Konno S, Miyamoto M, et al. Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ). Part 2. Endorsement of the alternative item. *J Orthop Sci* 2007;12:241–8.
3. Ware JE Jr. SF-36 health survey update. *Spine* 2000;25: 3130–9.
4. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
5. Hirabayashi K, Miyakawa J, Satomi K, Maruyama T, Wakano K. Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. *Spine* 1981;6:354–64.
6. Yonenobu K, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the Japanese Orthopaedic Association scoring system for evaluation of cervical compression myelopathy. *Spine* 2001;26:1890–5.

## Appendix

## Twenty-four items as candidates for JOACMEQ (temporary English version)

With regard to your health during the last week, please circle one response for each of the following questions that best applies. If your condition varies depending on the day or the time, circle the response that corresponds to your condition at its worst.

- C01 Can you fasten the front buttons of your blouse or shirt with both hands?  
1) possible without difficulty  
2) possible if I spend time  
3) impossible
- C02 Can you eat a meal with your dominant hand using a spoon or a fork?  
1) possible without difficulty  
2) possible if I spend time  
3) impossible
- C03 Can you raise your arm? (Answer for the weaker side.)  
1) I can raise it straight upward  
2) possible, although the elbow and/or wrist is a little flexed  
3) possible up to shoulder level  
4) impossible
- C04 Can you walk on a flat surface?  
1) possible without difficulty  
2) possible but slowly without any support  
3) possible only with the support of a handrail, a cane, or a walker  
4) possible but slowly even with support  
5) impossible
- C05 Can you stand on either leg without holding onto something? (or the need to support yourself)  
1) possible on both legs individually for more than ten seconds  
2) possible on either leg for more than ten seconds  
3) impossible with either leg
- C06 Do you have urinary incontinence?  
1) no  
2) when sneezing or straining  
3) when retaining urine over a period of more than 2 hours  
4) frequently  
5) always
- C07 How often do you go to the bathroom at night?  
1) rarely  
2) once or twice  
3) three times or more
- C08 Do you have a feeling of residual urine in your bladder after voiding?  
1) rarely  
2) sometimes  
3) most of the time
- C09 Can you initiate (start) your urine stream immediately when you want to void?  
1) most of the time  
2) sometimes  
3) usually not
- C10 While in the sitting position, can you look up at the ceiling by tilting your head upward?  
1) possible without difficulty  
2) possible to some degree  
3) impossible
- C11 Can you drink a glass of water without stopping despite the neck symptoms?  
1) possible without difficulty  
2) possible to some degree  
3) impossible
- C12 Can you look at your feet when you go down the stairs?  
1) possible without difficulty  
2) possible to some degree  
3) impossible
- C13 While in the sitting position, can you turn your head toward the person who is seated to the side but behind you and speak to that person while looking at his/her face?  
1) possible without difficulty  
2) possible to some degree  
3) impossible

## Appendix

## Twenty-four items as candidates for JOACMEQ (in Japanese)

最近1週間ぐらいを思い出して、設問ごとに、あなたの状態にもっとも近いものの番号に○をつけてください。

日や時間によって状態が変わる場合は、もっとも悪かったときのものをお答えください。

## 最近の1週間:

- 問1 ブラウスやワイシャツなどの前ボタンを両手を使ってかけることができますか。  
1) 不自由なくできる 2) 時間をかければできる 3) できない
- 問2 きき手でスプーンやフォークを使って食事ができますか。  
1) 不自由なくできる 2) 時間をかければできる 3) できない
- 問3 片手をあげることができますか。(左右の手のうち悪いほうで答えてください)  
1) まっすぐ上にあげることができる  
2) すこし手が曲がるが上にあげることができる  
3) 途中まで(肩の高さぐらいまで)ならあげることができる  
4) できない
- 問4 平らな場所を歩くことができますか。  
1) 不自由なく歩くことができる  
2) ゆっくりとならば歩くことができる  
3) 支持(手すり、杖、歩行器など)があれば、歩くことができる  
4) 支持を使ってもゆっくりとしか歩くことができない  
5) できない
- 問5 手で支えずに片足立ちができますか。  
1) 両足とも10秒数える間以上できる  
2) どちらかの足は10秒数えるまではできない  
3) どちらの足もほとんどできない
- 問6 おしっこ(尿)を漏らすことがありますか。  
1) まったくない  
2) くしゃみや気脹ったときに漏れる  
3) 2時間以上おしっこ(排尿)しないと漏れる  
4) しばしば漏れる 5) いつも漏れる
- 問7 夜中に、トイレ(おしっこ(排尿))に起きますか。  
1) ほとんど起きることはない  
2) 一晩に1、2回起きる  
3) 一晩に3回以上起きる
- 問8 おしっこ(排尿)の後も、尿の残った感じがありますか。  
1) ほとんどのときに無い  
2) あるときと無いときがある  
3) たいていのときにある
- 問9 便器の前で(便器に座って)、すぐにおしっこ(尿)が出ますか。  
1) ほとんどのときすぐに出る  
2) すぐに出るときとすぐには出ないときがある  
3) たいていのときすぐには出ない
- 問10 いすに腰掛けて、首だけを動かして、自分の真上の天井をみることができますか。  
1) 不自由なくできる 2) 無理をすればできる 3) できない
- 問11 コップの水を一気に飲み干すことができますか。  
1) 不自由なくできる 2) 無理をすればできる 3) できない
- 問12 階段を下りるときに、足元を見ることができますか。  
1) 不自由なくできる 2) 無理をすればできる 3) できない
- 問13 いすに座って、後ろの席に座った人の顔を見ながら話をすることができますか。  
1) 不自由なくできる 2) 無理をすればできる 3) できない

- QOL01 How is your present health condition?  
 1) excellent  
 2) very good  
 3) good  
 4) fair  
 5) poor
- QOL02 Do you have difficulty in climbing up the stairs?  
 1) I have great difficulty  
 2) I have some difficulty  
 3) I have no difficulty
- QOL03 Do you have difficulty with one of the following: bending forward, kneeling or stooping? If you have difficulty with one of them, how difficult is it?  
 1) I have great difficulty  
 2) I have some difficulty  
 3) I have no difficulty
- QOL04 Do you have difficulty walking more than 15 minutes?  
 1) I have great difficulty  
 2) I have some difficulty  
 3) I have no difficulty
- QOL05 Have you been unable to do your work or ordinary activities as well as you would like?  
 1) I have not been able to do them at all  
 2) I have been unable to do them most of the time  
 3) I have sometimes been unable to do them  
 4) I have been able to do them most of the time  
 5) I have always been able to do them
- QOL06 Has your work routine been hindered because of the pain?  
 1) greatly  
 2) moderately  
 3) slightly (somewhat)  
 4) little (minimally)  
 5) not at all
- QOL07 Have you felt discouraged and depressed?  
 1) always  
 2) frequently  
 3) sometimes  
 4) rarely  
 5) never
- QOL08 Do you feel exhausted?  
 1) always  
 2) frequently  
 3) sometimes  
 4) rarely  
 5) never
- QOL09 Have you felt happy?  
 1) always  
 2) almost always  
 3) sometimes  
 4) rarely  
 5) never
- QOL10 Do you think you are in decent health?  
 1) yes (I am healthy)  
 2) fairly (my health is better than average)  
 3) not very much (my health is average)  
 4) barely (my health is poor)  
 5) not at all (my health is very poor)
- QOL11 Do you feel your health will get worse?  
 1) very much so  
 2) a little bit at a time  
 3) sometimes yes and sometimes no  
 4) not very much  
 5) not at all

普段の生活についての質問です。最近の1週間のことを思い出してください。一番よくあてはまる1), 2), 3), …の番号に○をつけて下さい

#### 最近の1週間:

- 問1 あなたの現在の健康状態をお答えください。  
 1)最高によい 2)とてもよい 3)よい 4)あまりよくない 5)よくない
- 問2 あなたは、からだのぐあいが悪いことから、階段で上の階へ上ることをむずかしいと感じますか。  
 1)とてもむずかしいと感じる  
 2)少しむずかしいと感じる  
 3)まったくむずかしいとは感じない
- 問3 あなたは、からだのぐあいが悪いことから、体を前に曲げる・ひざまずく・かがむ動作をむずかしいと感じますか。  
 どれかひとつでもむずかしく感じる場合は「感じる」としてください。  
 1)とてもむずかしいと感じる  
 2)少しむずかしいと感じる  
 3)まったくむずかしいとは感じない
- 問4 あなたは、からだのぐあいが悪いことから、15分以上つづけて歩くことをむずかしいと感じますか。  
 1)とてもむずかしいと感じる  
 2)少しむずかしいと感じる  
 3)まったくむずかしいとは感じない
- 問5 あなたは、からだのぐあいが悪いことから、仕事や普段の活動が思ったほどできなかったことがありますか。  
 1)いつもできなかった  
 2)ほとんどいつもできなかった  
 3)ときどきできないことがあった  
 4)ほとんどいつもできた  
 5)いつもできた
- 問6 痛みのために、いつもの仕事はどのくらい妨げられましたか。  
 1)非常に妨げられた 2)かなり妨げられた 3)少し妨げられた  
 4)あまり妨げられなかった 5)まったく妨げられなかった
- 問7 あなたは落ち込んでゆううつな気分を感じましたか。  
 1)いつも感じた 2)ほとんどいつも感じた 3)ときどき感じた  
 4)ほとんど感じなかった 5)まったく感じなかった
- 問8 あなたは疲れ果てた感じでしたか。  
 1)いつも疲れ果てた感じだった  
 2)ほとんどいつも疲れ果てた感じだった  
 3)ときどき疲れ果てた感じだった  
 4)ほとんど疲れを感じなかった  
 5)まったく疲れを感じなかった
- 問9 あなたは楽しい気分でしたか。  
 1)いつも楽しい気分だった  
 2)ほとんどいつも楽しい気分だった  
 3)ときどき楽しい気分だった  
 4)ほとんど楽しくなかった  
 5)まったく楽しくなかった
- 問10 あなたは、自分は人並みに健康であると思いますか。  
 1)「人並みに健康である」と思う  
 2)ほぼ「人並みに健康である」と思う  
 3)かろうじて「人並みに健康である」と思う  
 4)「人並みに健康である」とはあまり思わない  
 5)「人並みに健康である」とはまったく思わない
- 問11 あなたは、自分の健康が悪くなるような気がしますか。  
 1)悪くなるような気が大いにする  
 2)悪くなるような気が少しする  
 3)悪くなるような気がするときもしないときもある  
 4)悪くなるような気はあまりしない  
 5)悪くなるような気はまったくしない

*Original article*

## Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ): part 4. Establishment of equations for severity scores

Subcommittee on Low Back Pain and Cervical Myelopathy, Evaluation of the Clinical Outcome Committee of the Japanese Orthopaedic Association

MITSURU FUKUI<sup>1</sup>, KAZUHIRO CHIBA<sup>2</sup>, MAMORU KAWAKAMI<sup>3</sup>, SHINICHI KIKUCHI<sup>4</sup>, SHINICHI KONNO<sup>4</sup>, MASABUMI MIYAMOTO<sup>5</sup>, ATSUSHI SEICHI<sup>6</sup>, TADASHI SHIMAMURA<sup>7</sup>, OSAMU SHIRADO<sup>8</sup>, TOSHIHIKO TAGUCHI<sup>9</sup>, KAZUHISA TAKAHASHI<sup>10</sup>, KATSUSHI TAKESHITA<sup>6</sup>, TOSHIKAZU TANI<sup>11</sup>, YOSHIAKI TOYAMA<sup>2</sup>, KAZUO YONENOBU<sup>12</sup>, EIJI WADA<sup>13</sup>, TAKASHI TANAKA<sup>14</sup>, and YOSHIO HIROTA<sup>15</sup>

<sup>1</sup>Laboratory of Statistics, Osaka City University Faculty of Medicine, Osaka, Japan

<sup>2</sup>Department of Orthopaedic Surgery, Keio University, 35 Shinanomachi, Shinjuku, Tokyo 160-8582, Japan

<sup>3</sup>Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan

<sup>4</sup>Department of Orthopaedic Surgery, School of Medicine, Fukushima Medical University, Fukushima, Japan

<sup>5</sup>Department of Orthopaedic Surgery, Nippon Medical School, Tokyo, Japan

<sup>6</sup>Department of Orthopaedic Surgery, The University of Tokyo, Tokyo, Japan

<sup>7</sup>Department of Orthopaedic Surgery, Iwate Medical University School of Medicine, Morioka, Japan

<sup>8</sup>Department of Orthopaedic Surgery, Saitama Medical School, Saitama, Japan

<sup>9</sup>Department of Orthopaedic Surgery, Yamaguchi University School of Medicine, Yamaguchi, Japan

<sup>10</sup>Department of Orthopaedic Surgery, Graduate School of Medicine, Chiba University, Chiba, Japan

<sup>11</sup>Department of Orthopaedics, Kochi Medical School, Kochi, Japan

<sup>12</sup>Department of Orthopaedic Surgery, Hoshigaoka Koseinenkin Hospital, Osaka, Japan

<sup>13</sup>National Hospital Organization, Osaka-Minami Medical Center, Osaka, Japan

<sup>14</sup>Department of Internal Medicine, Houai Hospital, Osaka, Japan

<sup>15</sup>Department of Public Health, Osaka City University Faculty of Medicine, Osaka, Japan

### Abstract

**Background.** To establish a patient-oriented outcome measure for cervical myelopathy, a subcommittee of the Japanese Orthopaedic Association (JOA) developed a new scoring system to evaluate the overall clinical status of patients, which could be completed by patients themselves. The subcommittee completed three large-scale studies to select and modify questions derived from various preexisting outcome measures including Short Form-36, and then finalized and validated the questionnaire, which comprised 24 questions.

**Methods.** The finalized questionnaire was administered to 369 patients with cervical myelopathy due to disc herniation, spondylosis, or ossification of posterior longitudinal ligament by randomly selected board-certified spine surgeons. Patients with different severities of myelopathy were included to insure accuracy and responsiveness of this questionnaire against patients' different neurological status.

**Results.** Data of 236 patients were employed and were subjected to rigorous statistical analyses. There was no question that was difficult to answer and distribution of answers for each question was not concentrated to one choice, indicating the appropriateness of all 24 questions. Results of factor anal-

ysis suggested that the 24 questions could be divided into five different factors or functional domains. The factors were defined as follows: factor 1, lower extremity function; factor 2, quality of life; factor 3, cervical spine function; factor 4, bladder function; and factor 5, upper extremity function. Finally, equations that would yield scores for the five factors were assembled. The score to be used to represent the degree of patients' disability or status in each domain can be calculated by multiplying prefixed numbers of selected answers to questions by preassigned coefficients. Coefficients were defined to make the minimum score 0 and the maximum score 100.

**Conclusions.** We have successfully established a questionnaire that is able to demonstrate the status of patients suffering cervical myelopathy from five different aspects represented by five intuitive numerical scores. The final issue to be confirmed is the responsiveness of this questionnaire to changes in patients' status after various surgical and nonsurgical treatments.

### Introduction

The Japanese Orthopaedic Association scoring system for the evaluation of cervical myelopathy (JOA score)

was first established by a committee of the JOA chaired by Hirabayashi.<sup>1</sup> Since then, this scoring system has been accepted universally in Japan as a tool to measure the outcomes of surgical and nonsurgical treatments for various cervical spinal disorders that cause cervical myelopathy. The JOA score first appeared in the English literature in 1980 when Hirabayashi published an article describing surgical results in patients with ossification of the posterior longitudinal ligament of the cervical spine (OPLL) that underwent expansive open-door laminoplasty.<sup>2</sup> The JOA score underwent the first revision in 1994 to reflect the deficits in shoulder and elbow functions, which are often caused by cervical root lesions, and this revised version was officially translated into English.<sup>3</sup> Various modified versions of the JOA score have also been introduced in the western countries.<sup>4-6</sup> The JOA score is a disease-specific and physician-oriented system that mainly assesses the neurological status of the patient and enables surgeons to compare the changes in the neurological status of the patient before and after certain treatments. Due to emerging needs to evaluate the impairments in patients' activities of daily living (ADL), which is related directly to their quality of life (QOL), various patient-oriented outcome measures, for example, the Short Form (SF)-36, have been developed and adopted into clinical practices in different medical fields.<sup>7</sup> To take up such needs, the JOA together with the Japanese Society of Spine Surgery and Related Research (JSSR), formerly called the Japanese Spine Research Society (JSRS), has appointed several

members of the Clinical Outcomes Committee of the JOA to organize a subcommittee, the aim of which is to develop a completely new patient-oriented scoring system for the evaluation of clinical results in patients with back pain and cervical myelopathy. The subcommittee decided to construct a self-rating questionnaire that could be filled out by patients themselves. Candidates of the questions to be included in the questionnaire were selected and modified from various preexisting outcome measures including the SF-36, the Rolland and Morris Disability Questionnaire, and the Oswestry Disability Index.<sup>8,9</sup> The subcommittee completed three large-scale studies to select and validate the questions that would ultimately become parts of the new JOA scoring system and as a result, the questionnaires including 25 and 24 questions for the evaluation of back pain and cervical myelopathy respectively were finalized.<sup>10-12</sup> In the present study, the cervical version of the finalized questionnaires was administered to patients with cervical myelopathy of different severity to insure the accuracy and responsiveness of this questionnaire and the obtained data were subjected to factor analysis in order to divide the 24 questions into different functional domains. The titles of the domains were designated according to the context of the questions in each domain. Finally, the equations that would yield scores based on the answers to the questions, selected by the patients, in the different domains were established. The scores would represent the degree of patient disability in the functional domains.

**Table 1.** Sex and age distribution of patients

Sex	Age groups (years)	Severity of myelopathy			NA	Total
		Mild	Moderate	Severe		
Male	20-29	0	0	1		1
	30-39	5	6	3		14
	40-49	5	5	8		18
	50-59	16	18	6	1	41
	60-69	24	26	9		59
	70-79	7	9	7		23
	80-89	3	1	2		6
	>90	0	0	0		0
	Total	60	65	36	1	162
Female	20-29	0	0	0		0
	30-39	2	1	0		3
	40-49	3	6	0		9
	50-59	4	8	2		14
	60-69	11	8	4		23
	70-79	5	8	7		20
	80-89	2	2	1		5
	>90	0	0	0		0
	Total	27	33	14		74
Total		87	98	50	1	236

**Table 2.** Distribution of scores for subdomains in original scoring system of Japanese Orthopaedic Association

Domain	Function	Score	Response number
Motor	Upper extremity: feeding		
	Impossible	0	5
	Severe	1	22
	Moderate	2	56
	Mild	3	92
Motor	Normal	4	61
	Upper extremity: shoulder and elbow function		
	Severe	-2	8
	Moderate	-1	14
	Mild	-0.5	45
Motor	Normal	0	166
	Unknown		3
	Lower extremity: gait		
	Impossible	0	9
		0.5	1
	Severe	1	26
		1.5	23
	Moderate	2	39
		2.5	20
	Mild	3	59
Sensory	Normal	4	59
	Upper extremity		
	Severe	0	9
		0.5	40
	Moderate	1	111
Sensory	Mild	1.5	60
	Normal	2	16
	Trunk		
	Severe	0	1
		0.5	5
Sensory	Moderate	1	39
	Mild	1.5	26
	Normal	2	165
	Unknown		
	Lower extremity		
Sensory	Severe	0	4
		0.5	24
	Moderate	1	69
	Mild	1.5	46
	Normal	2	93
Bladder	Urinary dysfunction		
	Severe	0	3
	Moderate	1	40
	Mild	2	55
	Normal	3	138
	Unknown		

n = 236

### Materials and methods

The subcommittee randomly chose 369 out of 829 board-certified spine surgeons who were registered in the JSSR database and asked them to participate in the present survey. Each surgeon was asked to administer the questionnaire to patients with cervical myelopathy due to disc herniation, spondylosis, or OPLL. The surgeons were required to include at least one patient each with mild, moderate, and severe myelopathy according to the discretion of each surgeon. Patients with (1) myelopathy due to nondegenerative diseases, such as

trauma, tumor, and rheumatoid arthritis; (2) other musculoskeletal diseases that would affect the evaluation of myelopathy; (3) difficulties filling the questionnaire due to their specific physical (e.g., defects or impairments in the limbs) or mental conditions (e.g., dementia, disorientation); and (4) a history of previous spinal surgery, were excluded from the analysis. Those who participated in our previous studies were also excluded. The surgeons were also asked to assess the neurological status of the patients using the original JOA scoring system. This study was approved by the Ethics Committee of the JSSR, and informed consent was obtained

**Table 3.** Distribution of answers for 24 questions

Question	Response				
	1	2	3	4	5
Q1-1	108 (45.8)	103 (43.6)	25 (10.6)		
Q1-2	171 (72.5)	56 (23.7)	9 (3.8)		
Q1-3	154 (65.3)	49 (20.8)	23 (9.7)	10 (4.2)	
Q1-4	107 (45.3)	82 (34.7)	24 (10.2)	15 (6.4)	8 (3.4)
Q1-5	107 (45.3)	72 (30.5)	57 (24.2)		
Q1-6	153 (64.8)	49 (20.8)	19 (8.1)	12 (5.1)	3 (1.3)
Q1-7	77 (32.6)	114 (48.3)	45 (19.1)		
Q1-8	127 (53.8)	83 (35.2)	26 (11.0)		
Q1-9	143 (60.6)	73 (30.9)	20 (8.5)		
Q1-10	120 (50.8)	86 (36.4)	30 (12.7)		
Q1-11	140 (59.3)	70 (29.7)	26 (11.0)		
Q1-12	161 (68.2)	53 (22.5)	22 (9.3)		
Q1-13	91 (38.6)	90 (38.1)	55 (23.3)		
Q2-1	6 (2.5)	14 (5.9)	87 (36.9)	100 (42.4)	29 (12.3)
Q2-2	48 (20.3)	113 (47.9)	75 (31.8)		
Q2-3	43 (18.2)	126 (53.4)	67 (28.4)		
Q2-4	46 (19.5)	98 (41.5)	92 (39.0)		
Q2-5	25 (10.6)	45 (19.1)	95 (40.3)	51 (21.6)	20 (8.5)
Q2-6	38 (16.1)	49 (20.8)	75 (31.8)	50 (21.2)	24 (10.2)
Q2-7	24 (10.2)	31 (13.1)	105 (44.5)	53 (22.5)	23 (9.7)
Q2-8	24 (10.2)	38 (16.1)	113 (47.9)	40 (16.9)	21 (8.9)
Q2-9	14 (5.9)	46 (19.5)	104 (44.1)	50 (21.2)	22 (9.3)
Q2-10	26 (11.0)	68 (28.8)	51 (21.6)	56 (23.7)	35 (14.8)
Q2-11	38 (16.1)	76 (32.2)	73 (30.9)	36 (15.3)	13 (5.5)

Numbers given in parentheses are percentages of the total response

from each subject. The completed questionnaires and the results of the original JOA scoring system were collected and sent to the independent central organization where biostatisticians compiled the results and input patient data into a spreadsheet. Rigorous statistical analyses including factor analysis were performed using SPSS software (Version 12, SPSS, Chicago, IL, USA).

## Results

Three hundred and sixty-nine patients were initially recruited for this survey. Among them, 106 patients were excluded because they had other musculoskeletal disorders that could affect the evaluation of cervical myelopathy. Most of them had nonspecific low-back pain without neurological symptoms ( $n = 70$ ) and mild knee joint pain due to osteoarthritis ( $n = 22$ ). Twenty-six patients gave no answers to one or more questions and were also excluded. One other patient was excluded because of the discretion of the surgeon in charge. The data of the remaining 236 patients were employed. The sex and age distributions of patients and the distributions of the scores for different subdomains in the original JOA scoring system are given in Tables 1 and 2. Because the majority of the excluded patients were those having other musculoskeletal diseases due to

spondylosis or osteoarthritis, the average age was significantly higher in the excluded patients than those that were included. The average scores for shoulder/elbow function and lower motor function in the JOA scoring system were significantly lower in the excluded patients than those for the included subjects.

The incidence of unanswered questions was less than 5%, indicating that there was no question that was difficult to answer. There also was no question for which the distribution of the answer was concentrated to one choice (Table 3), indicating the appropriateness of all 24 questions.

Results of factor analysis revealed that there were six common factors whose eigenvalues exceeded 1.0, which are thought to be the factor having significant contributions to the result. We decided to employ the first five factors because the cumulative contribution rate of the first five factors reached 60% and the contribution of the sixth factor was less than 5% (Table 4).

According to the calculation of the factor loadings after orthogonal rotation using the direct oblimin method with the Kaiser normalization, correlations among the 24 questions and the selected five factors were reexamined. When the maximum factor loading of a question exceeded 0.40, that question was supposed to be correlated with the factor. All 24 questions were judged to have correlation with at least one of the five

**Table 4.** Results of factor analysis

Factor	Eigenvalue	Contribution rate (%)	Cumulative contribution rate (%)
1	<b>8.86</b>	36.9	36.9
2	<b>2.02</b>	8.4	45.3
3	<b>1.56</b>	6.5	51.8
4	<b>1.36</b>	5.6	57.5
5	<b>1.20</b>	5.0	62.5
6	<b>1.08</b>	4.5	67.0
7	0.81	3.4	70.3
8	0.71	3.0	73.3
9	0.67	2.8	76.1
10	0.62	2.6	78.7
11	0.53	2.2	80.9
12	0.51	2.1	83.0
13	0.48	2.0	85.0
14	0.43	1.8	86.8
15	0.41	1.7	88.5
16	0.40	1.7	90.2
17	0.38	1.6	91.7
18	0.35	1.5	93.2
19	0.32	1.4	94.6
20	0.30	1.3	95.8
21	0.28	1.2	97.0
22	0.27	1.1	98.1
23	0.24	1.0	99.1
24	0.22	0.9	100.0

$n = 236$

Bold typeface indicates eigenvalues over 1.0

factors, except for Q1-4 and Q1-12. Q1-4 had relatively higher factor loadings for both factor 1 and 5 and this question was judged to be correlated with both factors. Q1-12 was first judged to be weakly correlated with factor 5 with a factor loading of 0.37; however, the committee decided to correlate this question also to factor 3 after reading the context of the question and given that the factor loading of 0.32 was also moderately high (Table 5).

According to the interpretations of the context of the questions that were divided into the five factors, each factor was categorized as follows; factor 1: lower extremity function; factor 2: quality of life; factor 3: cervical spine function; factor 4: bladder function; and factor 5: upper extremity function.

To establish an equation to calculate the individual score for each factor/domain that would intuitively indicate the status of a patient with regard to the designated function, the questions that had the maximum absolute factor loading value were used to calculate the score for the factor. For example, Q1-5 was used to calculate the score for factor 1 (lower extremity function), because the factor loading of this question for factor 1 was markedly larger (0.58) than those for the other four factors (-0.10, -0.14, -0.12, 0.11). As described above, Q1-4 and Q1-12 were used to calculate the scores for both factors 1 and 5, and 3 and 5, respectively (Table 5). The score was derived by multiplying the prefixed number of the

**Table 5.** Factor loading after orthogonal rotation

Question	Factors				
	1	2	3	4	5
Q1-5	<b>0.58</b>	-0.10	-0.14	-0.12	0.11
Q1-4	<b>0.55</b>	0.05	-0.08	-0.12	<b>0.40</b>
Q2-3	<b>-0.39</b>	0.25	0.16	0.18	-0.04
Q2-4	<b>-0.39</b>	0.24	0.10	0.33	-0.09
Q2-2	<b>-0.52</b>	0.20	0.13	0.24	-0.03
Q2-7	0.10	<b>0.79</b>	0.00	0.08	-0.01
Q2-8	0.14	<b>0.73</b>	0.06	0.12	-0.06
Q2-11	-0.01	<b>0.68</b>	-0.21	0.03	0.01
Q2-6	0.05	<b>0.55</b>	0.25	-0.06	-0.06
Q2-5	-0.27	<b>0.39</b>	0.19	0.04	-0.13
Q2-10	0.23	<b>-0.49</b>	0.08	-0.10	0.15
Q2-1	0.12	<b>-0.60</b>	-0.04	0.01	0.09
Q2-9	0.13	<b>-0.66</b>	-0.13	0.10	-0.06
Q1-11	-0.02	0.03	<b>-0.57</b>	-0.24	0.10
Q1-13	0.05	-0.06	<b>-0.75</b>	0.00	0.05
Q1-10	0.04	-0.01	<b>-0.87</b>	0.07	-0.06
Q1-12	0.10	0.04	-0.32	-0.21	<b>0.37</b>
Q1-9	0.08	-0.05	-0.07	<b>-0.43</b>	-0.01
Q1-7	0.14	-0.04	0.09	<b>-0.53</b>	-0.03
Q1-6	0.00	0.02	0.01	<b>-0.58</b>	0.03
Q1-8	-0.11	0.01	-0.06	<b>-0.61</b>	0.00
Q1-2	0.12	0.02	0.02	0.10	<b>0.73</b>
Q1-1	0.24	-0.13	0.01	0.02	<b>0.60</b>
Q1-3	-0.27	-0.11	-0.05	-0.11	<b>0.60</b>

Method of factor extraction: unweighted least-squares method. Orthogonal rotation: direct oblimin method with Kaiser normalization  
Bold typeface indicates absolute value of the factor loading of more than 0.35

**Table 6.** Coefficients for calculation of severity score

Question	Factors				
	1 Lower extremity function	2 Quality of life	3 Cervical spine function	4 Bladder function	5 Upper extremity function
Q1-1					-10
Q1-2					-15
Q1-3					-5
Q1-4	-10				-5
Q1-5	-10				
Q1-6				-10	
Q1-7				-5	
Q1-8				-10	
Q1-9				-5	
Q1-10			-20		
Q1-11			-10		
Q1-12			-5		
Q1-13			-15		-5
Q2-1		-3			
Q2-2	15				
Q2-3	5				
Q2-4	5				
Q2-5		2			
Q2-6		2			
Q2-7		5			
Q2-8		4			
Q2-9		-3			
Q2-10		-2			
Q2-11		3			

**Table 7.** Equations to calculate scores for different factors

Factor	Equation
1 Lower extremity function	$(Q1-4 \times 10 + Q1-5 \times 10 + Q2-2 \times 15 + Q2-3 \times 5 + Q2-4 \times 5 - 45) \times 100 + 105$
2 Quality of life	$(Q2-1 \times 3 + Q2-5 \times 2 + Q2-6 \times 2 + Q2-7 \times 5 + Q2-8 \times 4 + Q2-9 \times 3 + Q2-10 \times 2 + Q2-11 \times 3 - 24) \times 100 + 96$
3 Cervical spine function	$(Q1-10 \times 20 + Q1-11 \times 10 + Q1-12 \times 5 + Q1-13 \times 15 - 50)$
4 Bladder function	$(Q1-6 \times 10 + Q1-7 \times 5 + Q1-8 \times 10 + Q1-9 \times 5 - 30) \times 100 + 80$
5 Upper extremity function	$(Q1-1 \times 10 + Q1-2 \times 15 + Q1-3 \times 5 + Q1-4 \times 5 + Q1-12 \times 5 - 40) \times 100 + 95$

answer, selected by the patient, by the coefficient that were defined to make the difference between the minimum and maximum scores to be approximately 100 points (Table 6). The additional coefficients were also assigned to adjust the minimum score to be 0 and the maximum score to be 100. The final equations for the scores for the five domains are shown in Table 7.

## Discussion

In our previous studies, the questionnaires were constructed by referring to various preexisting outcome measures including the SF-36, the Roland and Morris Disability Questionnaire, and the Oswestry Disability Index,<sup>7-9</sup> which have commonly been used for patients

with different spinal disorders. We also took extra care to maintain relevance to the original JOA scoring systems by carefully assessing the correlation between the new and original systems. The new questionnaires have been revised several times and underwent many validation processes using rigorous statistical analyses. As a result, 24 questions were selected as the items of the finalized questionnaires for back pain and cervical myelopathy.<sup>10-13</sup>

In the present study, patients with different severity of cervical myelopathy were examined using the cervical version of the questionnaires to insure the accuracy and responsiveness of this questionnaire against various neurological states of the patient. Factor analysis was used to divide the 24 questions into different factorial domains, and the domains were categorized by inter-