Eur Spine J (2012) 21:130–137
DOI 10.1007/s00586-011-1992-2

ORIGINAL ARTICLE

Reliability and validity of the cross-culturally adapted French version of the Core Outcome Measures Index (COMI) in patients with low back pain

Stéphane Genevay · Christine Cedraschi · Marc Marty · Sylvie Rozenberg · Pierre De Goumoëns · Antonio Faundez · Federico Balagué · François Porchet · Anne F. Mannion

Received: 18 January 2011 / Revised: 22 June 2011 / Accepted: 16 August 2011 / Published online: 1 September 2011

Abstract

Purpose To conduct a cross-cultural adaptation of the Core Outcome Measures Index (COMI) into French according to established guidelines.

Methods Seventy outpatients with chronic low back pain were recruited from six spine centres in Switzerland and France. They completed the newly translated COMI, and the Roland Morris disability (RMQ), Dallas Pain (DPQ), adjectival pain rating scale, WHO Quality of Life, and EuroQoL-5D questionnaires. After ~14 days RMQ and COMI were completed again to assess reproducibility; a transition question (7-point Likert scale; “very much worse” through “no change” to “very much better”) indicated any change in status since the first questionnaire.

Results COMI whole scores displayed no floor effects and just 1.5% ceiling effects. The scores for the individual COMI items correlated with their corresponding full-length reference questionnaire with varying strengths of correlation (0.33–0.84, $P < 0.05$). COMI whole scores showed a very good correlation with the “multidimensional” DPQ global score (Rho = 0.71). 55 patients (79%) returned a second questionnaire with no/minimal change in their back status. The reproducibility of individual COMI 5-point items was good, with test–retest differences within one grade ranging from 89% for ‘social/work disability’ to 98% for ‘symptom-specific well-being’. The intraclass correlation coefficient for the COMI whole score was 0.85 (95% CI 0.76–0.91).

Conclusions In conclusion, the French version of this short, multidimensional questionnaire showed good psychometric properties, comparable to those reported for German and Spanish versions. The French COMI represents a valuable tool for future multicentre clinical studies and surgical registries (e.g. SSE Spine Tango) in French-speaking countries.

Electronic supplementary material The online version of this article (doi:10.1007/s00586-011-1992-2) contains supplementary material, which is available to authorized users.

S. Genevay (✉)
Division of Rheumatology, University Hospital of Geneva,
Hôpital Beau-Séjour, 26 av. Beau-Séjour, 1211 Geneva 14,
Switzerland
e-mail: stephane.genevay@hcuge.ch

C. Cedraschi
Division of General Medical Rehabilitation,
University Hospital of Geneva, Geneva, Switzerland

M. Marty
Division of Rheumatology, CHU Henri-Mondor Créteil,
Paris, France

S. Rozenberg
Division of Rheumatology, CHU de la Pitié-Salpetrière,
Paris, France

P. De Goumoëns
Spine Centre, Centre Hospitalier Universitaire Vaudois,
Lausanne, Switzerland

A. Faundez
Division of Orthopaedic Surgery,
University Hospital of Geneva, Geneva, Switzerland

F. Balagué
Division of Rheumatology, HRF-Hôpital Cantonal,
Fribourg, Switzerland

F. Porchet · A. F. Mannion
Schulthess Clinic, Zurich, Switzerland

Keywords Low back pain · Outcome measure · Questionnaire · Cross-cultural adaptation

Introduction

Common low back pain (CLBP) is an extremely frequent medical condition that can have important personal and societal repercussions. Considering the moderate effectiveness of most therapies, the need to combine several interventions to improve outcomes, and the high costs generated by patients with CLBP, formal evaluation may help clinicians in the implementation of their therapeutic interventions.

It has been recognized that patient-based outcome measures are the best way to evaluate patients with CLBP and it has been recommended that at least five domains should be explored: pain symptoms, LBP specific function, well-being, work disability and social disability, with patient satisfaction with care also being investigated following treatment [1]. For most of these domains (i.e. pain, function and general well-being) specific questionnaires have been developed and validated. Although potentially more precise, the combination of these questionnaires makes evaluations lengthy for patients and researchers. In addition, cumbersome evaluation tools are not feasible in daily clinical practice and this limits their implementation in large databases and usefulness for clinicians in decision-making at the patient level. To solve this problem a group of experts developed a short list of questions drawn from existing validated full-length questionnaires [1]. This short multidimensional questionnaire was originally introduces as "the core set" and more recently entitled "the Core Outcome Measure Index" (COMI) [2]. The questionnaire covers all the aforementioned domains, plus "general quality of life", each with one question. Thus, the COMI has the capacity to combine a multidimensional evaluation in a 7-question format.

Cross-cultural adaptations of the COMI exist for the German language [3] and for Spanish-speaking patients [4], and the COMI has become the main tool for the spine surgery registry of the Spine Society of Europe (Eurospine), Spine Tango [5]. In order to facilitate the widespread use of a questionnaire it is important to increase the number of validated language versions in which it is available [6]. This allows the questionnaire to be used in other countries, widens the language-based inclusion criteria for patients in clinical trials, and increases the number of studies available for meta-analysis.

The aims of this study were to conduct a cross-cultural adaptation of the COMI for use in French-speaking countries and to investigate the psychometric properties of the French version in patients with LBP consulting in different outpatient settings.

Materials and methods

The Core Outcome Measure Index

The domains included in the COMI are pain symptoms (two items separately for back and leg pain), function, symptom-specific well-being, and generic quality of life (QoL) (all in the past week), and work and social disability in the previous month. Pain scores are indicated on a 0–10 graphic rating scale. The response categories for the other items are 5-point adjectival or Likert scales. The two disability items have five response categories indicating the number of days with repercussions during the past month. A score for each subscale and an overall score are calculated. The pain score is given by the higher of the two pain-scale scores (back or leg). For the other items each incremental "step" is given 2.5 points so that they range from 0 (excellent condition) to 10 (worst condition). The scores for social disability and work disability are averaged to form one disability score. An overall score from 0 (best health status) to 10 (worst health status) can then be computed by the addition of the five subscales (pain, function, symptom-specific well-being, general QoL and disability) divided by 5.

Translation and cross-cultural adaptation

The translation and the cross-cultural adaptation of the original COMI English version into French were carried out in accordance with previously published guidelines [7] (see supplementary material).

Translation

Two native French speakers (T-1, T-2) carried out independent translations from English to French. The translators had different educational and job profiles. T-1 was a spine surgeon familiar with the concepts being examined, the clinical content of the questionnaires and with other disability questionnaires for LBP patients. T-2 had a degree in Fine Arts and was neither aware nor informed of the concepts being quantified and had no medical background (the ''naive translator'' [7]).

The different profiles of the two translators assured good agreement and accuracy with the original version in terms of both the content and the terminology. The two translations were compared with one another and with the English version. The two translators and a bilingual recording observer discussed any discrepancies until a consensus was reached. The results of the two translations were then synthesized into a common French translation, T-12.

Back-translation

Two native English speakers (one North American and one British) with French as their second-language (BT-1, BT-2) carried out two back-translations of the French version (BT-12) into English. Neither of the back-translators was familiar with the concepts explored. Both were blind to the English original and carried out their translation independently.

Expert committee

A committee was formed consisting of one of the translators, one of the back-translators, two clinicians (neurosurgeon and psychologist), and one methodologist/clinical research scientist (involved in the various translations/cross-cultural adaptations of the COMI). The translations, back-translations, and notes made in carrying out/comparing the translations were used to develop a ''pre-final'' version of the French COMI. The task of the committee was to guarantee semantic and conceptual equivalence between the French and English versions. Consensus was found for all parts of the questionnaire. All stages of the translation process were documented in written form.

Test of the pre-final version

A group of approximately 20 people (patients with back problems from the divisions of neurosurgery and orthopaedic surgery from one of the centres) were given the pre-final version of the French COMI questionnaire. After completion, a research assistant checked their responses and asked for their general comments (wording, ambiguities, ease of understanding, etc.). The findings from this phase of the adaptation process (face validity) were considered by the work-group when the final French version was produced.

Questionnaire battery

After informed consent was obtained, patients received a booklet of questionnaires. In addition to questions on demographics, LBP history and the French version of the COMI, this booklet contained validated translations of full-length questionnaires exploring most of the domains covered by the COMI (Table 1).

Patients

Seventy French-speaking patients with LBP were recruited from rheumatology or orthopedic outpatient spine clinics (Table 2). Inclusion criteria were: a low back problem causing back pain or referred pain for >3 months and ability to understand written French. Exclusion criteria were: specific LBP (i.e. fracture, cancer, infection or inflammatory disease), and the potential for a very rapid recovery (in the clinician's opinion). After giving informed consent, patients were required to complete the booklet and were then provided with a second shorter booklet to be completed and posted back 1 week later. This second booklet contained the COMI, the RMQ and a transition question evaluating any change in back status perceived since the time of the first booklet: 1 = very much worse, 2 = quite a bit worse, 3 = a bit worse, 4 = unchanged, 5 = a bit better, 6 = quite a bit better, 7 = very much better [8].

Of the 70 patients recruited, 67 (96%) returned a second questionnaire. Of these, 55 (79%) reported no or only minimal changes in their back pain status and this group was used for the analysis of test–retest reproducibility in patients with "stable" symptoms. Hence, the data of 70 patients (see Table 2 for patient characteristics) were used for the analyses of floor/ceiling effects and construct validity, and the data of 55 patients [33 women, 22 men; mean (SD) age 47 (16) years] were used for the assessment of reproducibility.

The study was approved by the local ethical committees.

Statistical analysis

Scores for each instrument were calculated as explained above, applying the following rules for missing data, as per the original questionnaires: no missing allowed for COMI and EQ-5D, since these have just one item per domain; for the DPQ 1 missing allowed for each domain; for WHOQoL-bref, a minimum of 80% answers were required for each domain [9].

Floor and ceiling effects were determined by calculating the number of individuals obtaining scores equivalent to the worst and the best status, respectively, for each item and for the global COMI score. This indicates the proportion of patients for whom no meaningful change in their condition could be detected (deterioration or improvement) as they are already at the extreme of the range. Floor/ceiling effects >70% are considered to be adverse and <15%, ideal [10].

Construct validity addresses the extent to which a questionnaire's scores relate in the expected manner to those of other instruments measuring a similar construct. The relationship was evaluated using Spearman Rank correlation coefficients, corrected for ties. Spearman's Rho coefficients were interpreted as follows: Rho $\geq$ 0.81–1.0 = excellent, 0.61–0.80 = very good, 0.41–0.60 = good, 0.21–0.40 = fair, and 0–0.20 = poor [11, 12]. Good to excellent coefficients were expected for the relationship between each item of the COMI and its reference

Table 1 Domains explored in the booklet and their corresponding questionnaire of reference

Domain	Questionnaire	Type of scales, number of items	Range of score
Pain	Adjectival pain scale	Likert, 1	1–5
Function	Roland & Morris Questionnaire (RMQ) [18][a]	Yes or no, 24	0–24
	Dallas Pain Questionnaire (DPQ) [13, 14], daily activities subscale[b]	Likert, 7	0–100
Quality of Life	World Health Organization Questionnaire (WHOQoL-bref)[c] [9]	Lickert, 26	16–80
Health related quality of life	European 5 Dimension Questionnaire (EQ-5D)[d] [19]	Likert, 5	−0.59 to 1
Social disability	Dallas Pain Questionnaire (DPQ) Social interest subscale	Likert, 3	0–100
Work disability	Dallas Pain Questionnaire (DPQ) Work–leisure subscale	Likert, 3	0–100
Mood	Dallas Pain Questionnaire (DPQ) Anxiety–depression subscale	Likert, 3	0–100
Multidimensional evaluation	Dallas Pain Questionnaire (DPQ) Whole score	Likert, 16	0–100

[a] RMQ: enquires as to whether back pain hinders the performance of 24 activities of daily living (today) with a score ranging from 0 (best health status) to 24 points (worst health status)

[b] DPQ: includes 16 items in four domains: daily living, work/leisure, anxiety/depression and social interest. Each scale goes from 0 (no problem) to 100 (fully disabled)

[c] WHOQoL-bref: measures four domains (each scored 4 (best status) to 20 (worst status)) considered to contribute to overall QoL: psychological, physical, social, and environmental well-being

[d] EQ-5D: evaluates mobility, self-care, usual activities, pain/discomfort and anxiety/depression, considering three levels of severity (no problems, some or moderate problems, and severe problems). A non-weighted approach was used to score the EQ5-D [21]

questionnaire (e.g. COMI function with RMQ and DPQ daily living). Fair to good coefficients were expected between the COMI whole score and other specific questionnaires. A good to very good correlation was expected between the COMI whole score and the global DPQ score.

Test–retest reproducibility indicates the extent to which the same results are obtained on repeated administration of the given instrument when no change is expected. For this analysis, patients answering anything other than “no change”, “a bit better” or “a bit worse” on the transition question were excluded. For the 5-point ordinal scales, reproducibility was assessed by examining the proportion of participants recording test–retest differences for each item within a reference value of ±1 category (where at least 90% was considered acceptable) [13].

For scales/items yielding approximately normally distributed values (pain scales, COMI whole score, RMQ), the differences in means for the repeated trials were examined using one-way repeated measures ANOVA, with determination of the intraclass correlation coefficients (ICCs; model $\mathrm{ICC}_{\mathrm{agreement}\ 2,1}$) and their 95% confidence intervals. ICCs greater than 0.7 in groups of at least 50 patients are generally considered to indicate acceptable reliability [10]. Standard errors of measurement $\mathrm{SEM}_{\mathrm{agreement}}$ were used to indicate the absolute measurement error (“agreement” [10]) and to calculate the minimum detectable change ($\mathrm{MDC}_{95\%}$) for the instruments, i.e. the degree of change required in an individual’s score to establish it (with a given level of confidence) as being a real change, over and above measurement error. At the 95% confidence level, this is defined as $1.96 \times \sqrt{2} \times \mathrm{SEM}$ which is equivalent to $2.77 \times \mathrm{SEM}$.

Results

Cross-cultural adaptation

The French version of the COMI is presented in the Appendix. Only few difficulties arose during its development: (a) translation of “interference with normal work (including both work outside the home and housework)”. From the back-translation, it appeared that ‘travail à la maison et à l’extérieur’ (the initial wording in French) was taken to mean outdoor activities, (e.g. gardening), as opposed to work done (e.g. one’s job) outside of the home. Thus, the term “activités habituelles” was used for “normal work”, “travail” for “work outside the home”, and “activités domestiques” for “housework”. (b) Translation of “satisfaction with your overall medical care in the hospital”. It appeared that the commonly used expression “qualité générale des soins” missed the emphasis on the

Table 2 Patient characteristics

Total number	70
Sex (male/female)	27/43
Age mean ± SD (range)	47.6 ± 15.2 (23–87)
Diagnostic category	
Non-specific LBP	35
Radiating pain, below knee	19
Radiating pain, not below knee	8
Radicular pain (±LBP)	8
LBP before this episode	
Yes	60
No	10
Duration of current episode (months)	
3–6	10
>6 and <18	18
>18	42
Normal work	
Retired	10
No paid work	4
On benefits	8
Employee	1
Professional	47
Length of current sick leave	
Not applicable	17
Not on sick leave	23
<7 weeks	2
7 weeks–3 mo	4
>3 and <6 mo	6
>6 and <18mo	9
>18	9
Educational level	
Obligatory	20
Higher education	28
Professional diploma	20
Type of work done for most of work-life	
Sedentary	18
Physical	29
Mixture of sedentary and physical	19
Missing	4

individual and on medical care rather than on other aspects of care. We hence used “l’ensemble de votre prise en charge médicale” to address this issue.

Missing data

Data were generally very complete for the 70 baseline questionnaires: missing answers were seen in up to 4/70 (5%) patients for the demographic/pain history questions; in up to 3 (4%) patients for the individual EQ-5D items and EQ-5D whole score; in 6 (9%) patients for the EQ-5D general health VAS; and in up to 2 (3%) patients for the individual COMI items and COMI whole score. For the RMQ, 3/70 (4%) had too many missing answers to allow valid calculation of a score. For the WHOQol domains, a score could not be calculated for 7 (10%) patients for WHOQol physical, up to 16 (23%) for WHOQoL social, with 7 (10%) patients also having too many missing answers to calculate a WHOQoL total score. All scores could be calculated for the DPQ individual domains and whole score.

Floor and ceiling effects

The floor (worst status) and ceiling (best status) effects of the questionnaires are shown in Table 3.

Acceptably low floor effects were found for pain and QoL (3–6%) and for function (13.0%) but rather high values were found for both social and work disability (each 38%) and symptom-specific QoL (44%). A low floor effect was found for the COMI whole score (1.5%**).** A minimal ceiling effect was found for some of the individual COMI items (back pain, function and symptom-specific well-being) and for its whole score.

The EQ-5D showed generally low floor effects (3–6%) except for pain (24%) but ceiling effects were rather high (21–93%) for all domains other than pain. There were no floor or ceiling effects for the EQ-5D whole score or EQ-VAS general health status.

There were minimal floor (0–3%) and ceiling (0–7%) effects for the RMQ, the DPQ domains and whole score, and the WHOQoL domains and whole score.

Construct validity

The correlation coefficients for the relationship between the scores for each item of the COMI and its corresponding full-length questionnaire are shown in Table 4. An excellent correlation was found between the COMI pain score (i.e. the worst rating between back and leg pain numeric rating scale) and the adjectival pain-scale scores (Rho = 0.84). Good to very good correlations (Rho between 0.54 and 0.67) were found between the scores for the COMI function item and the full-length function/disability questionnaires (RMQ, DPQ daily living, WHOQoL physical). Neither of the full-length QoL questionnaires was found to have more than a fair correlation (Rho $\leq$ 0.43) with the scores for COMI symptom-specific well-being. The scores for the COMI general QoL item showed good to very good correlations (Rho between 0.54 and 0.67) with the scores for the global QoL scales. COMI social disability scores showed a fair correlation with DPQ social interest scale (Rho = 0.33), whilst COMI work disability scores correlated well (Rho = 0.55) with DPQ work/leisure scores.

Table 3 Floor and ceiling effects for all the instruments in the first evaluation

Instrument	Floor (worst status)[a]	Ceiling (best status)[a]
COMI low back pain	4.4	0
COMI leg pain	3.0	16.4
COMI worst pain (leg or back)	5.9	0
COMI function	13.0	0
COMI symptom-specific well-being	43.5	0
COMI quality of life	5.8	1.4
COMI social disability	37.7	11.6
COMI work disability	37.7	27.5
COMI whole score	*1.5*	*0*
Roland Morris score	*1.5*	*0*
EQ-5D mobility	2.9	52.2
EQ-5D self-care	0	92.5
EQ-5D usual activities	5.9	20.6
EQ-5D pain	23.9	0
EQ-5D anxiety/depression	5.9	32.4
EQ-5D total score	*0*	*0*
EQ-5D VAS general health	0	0
Dallas daily activities	*1.4*	*0*
Dallas work and leisure	*2.9*	*0*
Dallas depression/anxiety	*0*	*2.9*
Dallas social activities	*1.4*	*5.7*
Dallas whole	*0*	*0*
WHOQoL physical	*0*	*0*
WHOQoL psychological	*0*	*0*
WHOQoL social	*0*	*7.4*
WHOQoL environmental	*0*	*1.6*
WHOQoL whole score	*0*	*0*

Italicized rows indicate scores from scales with more than one item

[a] Floor/ceiling effects >70% are considered to be adverse and <15%, ideal

The correlation between the COMI whole score and the DPQ whole score was very good (Rho = 0.71). No items of the COMI had more than a good correlation with DPQ depression-anxiety (Rho between 0.30 and 0.54).

Test–retest reproducibility

The mean duration between the first and the second questionnaire was 14 (SD = 11) days.

Differences in response to each domain on the COMI were ±1 category in 50/54 (92.6%) patients for ‘function’, 53/54 (98.1%) for ‘symptom-specific well-being’, 54/55 (98.2%) for ‘general QoL’, 49/55 (89.1%) for ‘social disability’ and 49/55 (89.1%) for ‘work disability’. Only the disability items fell just short of the expected 90% level [13].

No systematic biases, that is, no significant differences in the mean values, for the repeated (test–retest) scores were found for any of the instruments with the exception of the COMI worst pain, which showed a slightly but significantly lower value in the retest, $P = 0.045$ (Table 5).

The ICCs were between 0.83 (COMI worst pain) and 0.87 (RMQ) indicating very good reproducibility (Table 5). The SEM and MDC95% values are shown in Table 5; expressed as a percentage of the maximum score range for the given scale, the SEMs were similar for all scales, being approximately 7–11%.

Discussion

The results of this study indicate that the French version of the COMI has acceptable psychometric properties. The cross-cultural adaptation of the English COMI was carried out following established guidelines [7], in an attempt to produce a reliable and valid adaptation of the questionnaire. No real problems were encountered in this process, except for ensuring the idiomatic equivalence of the question regarding interference of the back problem with normal work activities.

The most important methodological limitation resides in the high rate of floor and/or ceiling effects for two items, symptom-specific QoL and disability. The same problem was observed for these items in the German [3] and Spanish [4] versions of the COMI, but these studies also found high ceiling effects for function which was not the case in this study. This may be explained by differences in the patient groups as both German and Spanish studies were conducted mainly in patients referred for surgery, with generally more severe symptoms. Floor and ceiling effects are often encountered when the number of response categories is low, as exemplified in this study by some of the results for the EQ-5D. With only three response categories, some of the items had a ceiling effect up to 93%. Theoretically high floor or ceiling effects could influence the questionnaire’s responsiveness because for patients in these extremes, deterioration (or improvement respectively) cannot be measured. Responsiveness was not explored in this study; however, both the German and the Spanish versions have been shown to display excellent responsiveness [3, 4, 14]. Overall, the French version demonstrated good construct validity, comparable to the Spanish and German versions [3, 4]. Most importantly, for the first time the relationship between the COMI whole score and another multidimensional questionnaire, the DPQ, was examined, and found to be very good. The DPQ is the only available LBP questionnaire specifically designed in accordance with the biopsychosocial model of back pain [15]. Symptom-specific QoL was the only item

Table 4 Most relevant correlations between COMI subscales and full-length questionnaires

Other questionnaires	COMI							
	Pain (worst of back or leg)	Function	Symptom specific well-being	Generic quality of life	Social disability	Work disability	Average social and work disability	Whole index score
Pain								
Pain adjectival scale	0.84							
Function								
RMQ		0.54						0.62
DPQ: daily living		0.67						0.74
Quality of life								
WHOQOL-BREF			−0.43	−0.67				−0.59
Health state								
EQ-5D			−0.36	−0.54				−0.62
Social disability								
DPQ: social interest					0.33			0.46
Work disability								
DPQ: work-leisure						0.55		0.72
Average (work and social) disability								
DPQ: physical (mean daily living and work-leisure)							0.58	–
Mood								
DPQ: depression-anxiety	0.52	0.47	0.34	0.54	0.30	0.34	0.33	0.57
Multidimensional								
DPQ: global								0.71

Correlations in which Rho is greater than 0.3 are all significant at $P < 0.015$

Table 5 Test–retest reliability results for each of the domain index-items and the full reference scales

Instrument	No items	Range	M1	M2	P	ICC	95% CI_{ICC}	$SEM_{agreement}$	SEM%	MDC95%
COMI whole score	5	0–10	6.1 (1.8)	6.0 (1.9)	0.41	0.85	0.76–0.91	0.72	7.2	1.98
COMI back pain	1	0–10	5.4 (2.3)	5.2 (2.1)	0.17	0.85	0.76–0.91	0.84	8.4	2.32
COMI leg pain	1	0–10	4.1 (3.0)	3.8 (2.7)	0.25	0.85	0.75–0.91	1.10	11.0	3.04
COMI worst pain	1	0–10	6.0 (2.2)	5.6 (1.9)	0.05	0.83	0.73–0.90	0.83	8.3	2.31
Roland Morris disability	24	0–24	11.6 (5.1)	11.2 (5.3)	0.27	0.87	0.79–0.93	1.85	7.7	5.11

M1, M2 mean value at first and second assessment. *P* significance of difference between mean values on the two occasions (one way ANOVA with repeated measures), *ICC* intraclass correlation coefficient (ICC2.1), CI_{ICC} 95% confidence intervals for the ICC, *SEM* standard error of measurement, *SEM%* SEM as percentage of maximum score, *MDC95%* minimum detectable change score

that showed a low correlation with the corresponding full-length questionnaires. The same observation was also made with the German [3] and the Spanish versions [4]. We hypothesized that this item might better correlate with health-related QoL rather than QoL. To our surprise the opposite was found as the correlation with WHOQoL-bref was slightly superior to the correlation with EQ-5D. This item appears to explore a unique dimension, not captured by other questionnaires and which may warrant further studies.

As reported for the other language versions [3, 4], the COMI summary score showed very good reproducibility, in the same range as the RMQ. Two of the individual items (measuring disability) using 5-point Likert scale fell just short (89.1%) of the ideal 90% level for the proportion of test–retest differences ±1 category [13] and the item "worst pain" showed a slight systematic bias in the test–retest (values slightly lower in the second assessment). A similar trend was observed with the Italian version (A. Mannion, personal communication) but it did not influence the psychometric properties of the overall summary score. In this French version, the minimal detectable change ($MDC_{95\%}$, Table 4) for the COMI whole score computed from the test–retest was 1.98, which was fairly comparable to that found for the German version (1.74).

Conclusion

The French version of the COMI has acceptable psychometric properties. As such it can be considered to be a suitable instrument for implementation in the Spine Tango Registry or in any other multi-language databases of outcomes in LBP patients. Short, time-saving, easily scored, multi-dimensional questionnaires have been advocated as the best way to monitor patients with chronic musculoskeletal conditions [16]. The effectiveness of such methods in daily practice has been demonstrated for rheumatoid arthritis [17], for patients with inguinal hernia [18] and for back pain patients using the German version of the COMI [2]. We believe that the systematic and widespread use of the French version in similar settings might enhance the quality of the follow-up of patients with chronic LBP.

Conflict of interest None.

References

1. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G (1998) Outcome measures for low back pain research. A proposal for standardized use. Spine 23:2003–2013
2. Mannion AF, Porchet F, Kleinstuck FS, Lattig F, Jeszenszky D, Bartanusz V, Dvorak J, Grob D (2009) The quality of spine surgery from the patient’s perspective. Part 1: the Core Outcome Measures Index in clinical practice. Eur Spine J 18 Suppl 3: 367–373
3. Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Semmer NK, Jacobshagen N, Dvorak J, Boos N (2005) Outcome assessment in low back pain: how low can you go? Eur Spine J 14: 1014–1026
4. Ferrer M, Pellise F, Escudero O, Alvarez L, Pont A, Alonso J, Deyo R (2006) Validation of a minimum outcome core set in the evaluation of patients with back pain. Spine 31:1372–1379 (discussion 1380)
5. Zweig T, Mannion AF, Grob D, Melloh M, Munting E, Tuschel A, Aebi M, Roder C (2009) How to Tango: a manual for implementing Spine Tango. Eur Spine J 18 Suppl 3:312–320
6. Costa LO, Maher CG, Latimer J (2007) Self-report outcome measures for low back pain: searching for international cross-cultural adaptations. Spine 32:1028–1037
7. Beaton DE, Bombardier C, Guillemin F, Ferraz MB (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 25:3186–3191
8. Beurskens AJ, de Vet HC, Koke AJ (1996) Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 65:71–76
9. Health PoM (1996) WHOQOL bref introduction, administration scoring and generic version of the assesment. In: WHO (ed) WHO, Geneva
10. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34–42
11. Steiner D, Norman G (1995) Health measurement scales: a practical guide to their development and use. Oxford Medical Publications, Oxford
12. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum Associates, Hillsdale NJ
13. Nevill AM, Lane AM, Kilgour LJ, Bowes N, Whyte GP (2001) Stability of psychometric questionnaires. J Sports Sci 19:273–278
14. Marty M, Blotman F, Avouac B, Rozenberg S, Valat JP (1998) Validation of the French version of the Dallas Pain Questionnaire in chronic low back pain patients. Rev Rhum Engl Ed 65: 126–134
15. Lawlis GF, Cuencas R, Selby D, McCoy CE (1989) The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior. Spine (Phila Pa 1976) 14: 511–516
16. Pincus T, Yazici Y, Sokka T (2007) Quantitative measures of rheumatic diseases for clinical research versus standard clinical care: differences, advantages and limitations. Best Pract Res Clin Rheumatol 21:601–628
17. Fransen J, Moens HB, Speyer I, van Riel PL (2005) Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial. Ann Rheum Dis 64:1294–1298
18. Staerkle RF, Villiger P (2011) Simple questionnaire for assessing core outcomes in inguinal hernia repair. Br J Surg 98:148–155
19. Brooks R (1996) EuroQol: the current state of play. Health Policy 37(1):53–72