Reliability and validity study on the Hungarian versions of the Oswestry Disability Index and the Quebec Back Pain Disability Scale

Tamás Valasek · Peter Paul Varga ·
Zsolt Szövérfi · Michelle Kümin ·
Jeremy Fairbank · Aron Lazary

Received: 15 January 2012 / Revised: 21 November 2012 / Accepted: 22 December 2012 / Published online: 16 January 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract

Purpose Patient-reported outcome measurements (PROMs) are widely used in spine care. The development of reliable and valid National versions of spine-related disability questionnaires is strongly recommended from both the clinical and scientific points-of-view. The aims of this study were to adapt and validate the Oswestry Disability Index (ODI) and the Quebec Back Pain Disability Scale (QDS) for use with the Hungarian language.

Methods After translating and culturally adapting the ODI and QDS, 133 patients with lumbar degenerative spinal disorder filled in the questionnaire booklet twice within 2 weeks. Subjects completed the Hungarian versions of the two PROMs as well as the WHOQoL-BREF validated as a general life quality questionnaire and Visual Analogue Scale of pain. Internal consistency, reliability and construct validity of the questionnaires

were determined, as were the standard error of measurement (SEM) and minimal detectable change (MDC) scores.

Results The Hungarian ODI consisted of one factor that showed good internal consistency (Cronbach- α 0.890). The QDS showed a four-factor structure with Cronbach- α values between 0.788 and 0.917. No significant floor or ceiling effects were observed. The test–retest analysis showed excellent reliability of the Hungarian ODI and QDS. The intraclass correlation coefficients (ICC) were 0.927 and 0.923, respectively. SEM values of 4.8 and 5.2 resulted in a MDC of 13 and 14 points in the Hungarian ODI and QDS, respectively. The correlation coefficient (r) between pain and ODI was 0.680 ($p < 0.001$) and the correlation between the ODI and the physical subscale of WHOQoL was also very good ($r = -0.705$, $p < 0.001$). The QDS total score and its four subscales correlated significantly with pain and with the physical subscale of WHOQoL ($r > 0.4$, $p < 0.001$). The level of disability measured by the Hungarian ODI and QDS was significantly higher in the surgical subgroup than in non-surgically treated patients ($p < 0.001$).

Conclusions Translation and cultural adaptation of the ODI and QDS were successful. Hungarian versions of the ODI and QDS proved to be reliable, valid PROMs confirming that they can be used in future clinical and scientific work with Hungarian-speaking spine patients.

Electronic supplementary material The online version of this article (doi:10.1007/s00586-012-2645-9) contains supplementary material, which is available to authorized users.

T. Valasek
Faculty of Health Sciences, Semmelweis University,
Vas U. 17, Budapest 1444, Hungary

P. P. Varga · Z. Szövérfi · A. Lazary (✉)
National Center for Spinal Disorders, Buda Health Center,
Királyhagyó U. 1-3, Budapest 1126, Hungary
e-mail: aron.lazary@bhc.hu

M. Kümin
Department of Physiology, Anatomy and Genetics,
University of Oxford, Le Gros Clark Building,
South Parks Rd, Oxford OX1 3QX, UK

J. Fairbank
Nuffield Orthopaedic Centre, Windmill Rd,
Headington, Oxford, Oxfordshire OX3 7HE, UK

Keywords Low back pain · Oswestry Disability Index · Quebec Back Pain Disability Scale · Hungarian version · Cross-cultural adaptation · Validity

Introduction

The need for patient-reported outcome measurements (PROMs) in orthopaedic procedures has been widely

reported in recent years. Numerous high-quality studies have previously shown the reliability and validity of different types of PROMs. The orthopaedic community has also stepped up its adoption of the use of questionnaires [1] emphasising the importance of cross-cultural adaptation and validation of these measurement tools [2–4]. Reliable and valid national versions of PROMs enable national and international condition-specific registries, with high levels of evidence, to be established [5, 6]. In spinal diseases, one of the major condition-specific outcomes is the level of disability. Spine-related disability originates from neurological deterioration and, more frequently, pain. Low back pain (LBP) and radicular leg pain (LP) are amongst the most common complaints of adults seeking medical help; they cause serious economical and social problems [7–9]. Recently published Hungarian epidemiological data showed the outstanding impact of LBP in a typical central European country [10]. Results of a clinical and radiological survey of 10,000 participants showed that high prevalence of LBP significantly increases with age. In those between 50 and 59 years old, the prevalence of LBP in the preceding month was 56.7 %. The lifetime prevalence of irradiating leg pain was 29.6 % in the whole cohort. More than a fifth of the participants (21.5 %) had previously been on sick-leave due to LBP and 3 % of the subjects had had previous spine surgery.

Several different PROMs, measuring spine-related disability, have been developed and published [3, 11]. The Oswestry Disability Index (ODI) [12, 13] and the Quebec Back Pain Disability Scale (QDS) [14, 15] are two reliable, valid and widely used PROMs in spine care that have been validated for use with numerous languages since first being reported in English. Validation studies of the Spanish [16], French-Canadian [17], Danish [18], Chinese [19, 20], German [21], Korean [22], Polish [23], Italian [24], Persian [25], Finnish [26], Turkish [27], Brazilian-Portuguese [28], Norwegian [29] and Greek [30] versions of the ODI have all been published. Nine validation studies on national versions of the QDS (Greek [31], Turkish [32], Polish [23], Persian [25], Portuguese [33], Korean [34], Dutch [35], French [36] and Arabic [37]) have been reported so far. All the papers above, in addition to comparison studies of the original English questionnaires [38, 39], reported these PROMs as reliable and valid measurement tools.

The epidemiological data above explicitly demonstrated the health care and social impact of spinal conditions as well as underlining the need for validated national versions of spine related PROMs. The aim of the present study was to cross-culturally adapt to the Hungarian language and to validate two spine-related disability questionnaires (ODI, QDS).

Materials and methods

Subjects

Between March and June 2011, 170 patients diagnosed consecutively with lumbar degenerative spinal disorder at the outpatient clinic of the National Center for Spinal Disorder (Budapest, Hungary) were asked to complete a questionnaire booklet and were screened for inclusion in the study (first test). Based on clinical data, patients with serious neurological conditions requiring acute surgery were excluded. One hundred and fifty patients were selected to participate in the study according to the following criteria: native Hungarian; of adult age (≥ 18); LBP and/or sciatica (LP) for more than 3 months; able to read and answer questionnaires in Hungarian. All subjects signed a written consent form describing the scientific purpose of the systematic collection and use of their clinical data. Consented patients who were scheduled for a second visit to the clinic within 2 weeks of the first test (because of a physiotherapy session or anaesthesiologist/surgeon check-up, etc.) were asked to complete the questionnaires again (retest). This study was approved by the Ethics Committee of the Hungarian Ministry of Health.

Measurement tools

Subjects completed the ODI (version 2.1a) and the QDS as condition-specific PROMs, the general construct of which have been published previously several times [11, 40]. The total score of both PROMs were converted to rate of disability in percentage using the proportional recalculation method. According to the developers' instructions, the ODI score was only calculated if nine out of ten questions had been answered. The QDS score (and the scores of the QDS subscales) was calculated if more than 50 % of the 20 questions had been answered. The proportional recalculation method means that the raw sum score is divided with the possible maximum score (after taking into consideration the missing answers) and multiplied by 100.

The questionnaire booklet also contained a 10-cm long visual analogue scale for pain and the previously validated Hungarian version of the WHOQoL-BREF (WHOQoL) general life quality questionnaire [41]. The WHOQoL contains 26 items about different aspects of life. The first two questions have to be analysed separately as 5-point Likert scales. The other 24 items constitute four domains (physical, psychological, social relationships and environmental domains) which are scored in a positive direction (i.e. higher scores denote a higher quality of life) and transformed to a 0–100 scale using the formula described by the WHOQoL user manual.

Translation and cross-cultural adaptation

Translation and cross-cultural adaptation of the ODI and QDS were both carried out by following the guidelines and instructions of the MAPI Research Institute (the official distributor of the licensed questionnaires) [42]. Firstly, the concept investigated by each item of the two original PROMs was clarified and discussed with the developers of the original instruments. Secondly, a three-step translation and cross-cultural adaptation was performed. Forward translation: two professional native Hungarian translators translated the original (English) questionnaires independently into Hungarian. One of the translators specialises in medical terminology. Following a reconciliation meeting between the study investigators and the forward translators, a consensus for the first Hungarian version (V1) was reached. Backward translation of V1 was performed by a third, independent, professional translator of native English origin (British) who was bilingual in English and Hungarian. The backward-translated versions of the ODI and QDS were then compared with the original versions by an expert committee comprising one of the forward translators, the backward translator, two spine-surgeons, a rheumatologist, a physiotherapist, a psychologist and a patient-representative. The expert committee checked that the conceptual content of the forward consensus versions remained. This created the second Hungarian versions of the two PROMs (V2). The V2 versions were then tested on a heterogeneous group of eight Hungarian patients suffering from low-back pain (carefully selected on the basis of age and different educational levels). The patients could discuss any general or specific questions with the investigator controlling the pilot procedure. Upon completion, the subjects were briefly interviewed about the meaning of each item. In parallel, ten further clinicians (spine surgeons and physiotherapists) were asked for a clinician's review of the two V2 versions. Results of the pilot testing were taken into consideration when producing the final versions of the Hungarian ODI and the QDS (V3) used in this study.

Statistical analysis

Internal homogeneity of the Hungarian versions of the two PROMs was tested by calculating the Cronbach- α value, taking into consideration the factor structure of the questionnaires. Exploratory factor analysis (EFA) was applied to determine the factor structure of the Hungarian ODI and QDS. Kaiser–Meyer–Olkin's measure of sample adequacy (KMO) and Bartlett's test of sphericity were applied to test the sample and the appropriateness of the factor analysis. Principal component analysis as an extraction method with varimax rotation was used to explore the factor structure. Because the number of variables was less than 30 and the

mean of the communalities after extraction was >0.7 , we used the Kaiser criterion to determine the factors with eigenvalues more than 1. Floor and ceiling effects were determined using not only the end-range scores of the original questionnaires but also the minimum score plus and the maximum score minus the “minimal detectable change” (MDC). Construct validity was demonstrated by determining the difference in disability between the surgical and non-surgical patient groups as well as the changes in the ODI and the QDS scores in the presence of nerve root deficit (motor weakness and/or sensory disturbance) (t tests). We hypothesised that the level of disability would be higher in patients requiring spinal surgery and with nerve root deficit. Associations between pain, WHOQoL subscales and the disability scores were also analysed calculating the Pearson's correlation coefficients where $r > 0.40$ was considered satisfactory ($r > 0.80$ as excellent, 0.61–0.80 very good, 0.41–0.60 good, 0.21–0.40 fair, 0–0.20 poor). Based on the concept of the ODI and the QDS, the pre-defined hypotheses were that pain and the physical subscale of WHOQoL were expected to correlate with the ODI and the QDS (sub)scales and that the correlation between disability and the psychological, social and environmental subscales of WHOQoL would be weak ($r = <0.4$). The results of the baseline measurement (test) were used in these analyses. To examine reproducibility, we carried out test–retest analyses. Differences in the mean values of the ODI and the QDS for the test–retest measurements were analyzed with paired t tests. Reliability was demonstrated by calculating the intraclass correlation coefficients (ICC) for absolute agreement in a two-way random effects ANOVA model. Agreement was demonstrated by determining the standard error of measurement (SEM) for the repeated measurements. The SEM was used to calculate the minimal detectable change at 95 % confidence level ($MDC_{95\%}$) for the ODI and the QDS applying the $2.77 \times SEM$ formula [43]. All statistical analyses were performed with SPSS 15.0 software; p values of less than 0.05 were considered significant.

Results

Study cohort

One hundred and fifty patients were included in the study. Nine of them missed the second visit and, consequently, did not fill in the retest study questionnaires within 2 weeks. Eight more subjects were excluded from further analyses because of more than one missing answer in the ODI. Thus, the final study cohort consisted of 133 subjects (Table 1). For the majority of patients ($N = 112$), spinal surgery was indicated. The remaining 21 patients

Table 1 Study cohort—baseline data ($N = 133$)

Male/female; N (%)	56/77 (42/58)
Age; years, mean (SD)	48.1 (15.3)
Weight; kg, mean (SD)	79.1 (15.9)
Height; cm, mean (SD)	170.7 (8.9)
BMI; kg/m^2 , mean (SD)	27.1 (5.0)
Pain; VAS, mean (SD)	6.8 (2.0)
Duration of symptoms; months, N (%)	
4–6	46 (34.6)
7–12	32 (24.1)
13–24	16 (12.0)
>24	39 (29.3)
Diagnosis; N (%)	
Lumbar disc herniation	57 (42.9)
Mechanical low back pain	27 (20.3)
Lumbar spinal stenosis	26 (19.5)
Failed back surgery syndrome	23 (17.3)
Primary pain source; N (%)	
Low back pain	73 (54.9)
Lower extremity pain	60 (45.1)
Nerve root deficit; N (%)	60 (45.1)

comprised a non-surgical subgroup. Seventy-three patients (55 %) had no nerve root deficit others while 60 patients (45.1 %) had lower extremity paresis and/or sensory disturbance.

Cross-cultural adaptation

The Hungarian versions of the ODI (version 2.1a) and the QDS are shown in the Supplement Material. During the pilot study, interpretation difficulties arose only in relation to the Hungarian version of the ODI. The fourth and fifth answers to section 10 (“Traveling”) caused interpretation problems. Even though these sentences were accurately translated, patients still had difficulty in understanding them. The original version uses the verb “to restrict” (fourth “Pain restricts me to journey of <1 h.” fifth “Pain restricts me to short necessary journey under 30 min.”). In V2, we used one of the possible Hungarian translations of this verb (“akadályoz”) but some of the pilot patients had difficulty in interpreting it. In the final version, another synonym was used (“korlátoz”) with a modified word order. These alterations brought about no changes in meaning but made it easier for the patients to interpret the sentences. According to other national versions, in item 4 of the ODI, we used metric equivalents of the English version’s 1 mile, quarter of a mile and 100 yards. Instead of the strict metric equivalents—1.6 km, 402 m, 91 m—we used 1.5 km, 400 m and 100 m without any interpretation problem.

Missing data, internal consistency and factor structure, floor and ceiling effect

In the final cohort ($N = 133$), two patients failed to answer the “personal care” question (ODI-2) and 44 patients omitted the “sex life” section (ODI-8). The highest ratio of missing answers in the QDS occurred in category 11 (“throw a ball”) which was omitted by six patients (Table 2). The KMO value was high (0.85 and 0.90 in the ODI and QDS, respectively) and the Bartlett’s test of sphericity was significant (ODI: $\chi^2 = 320.2$, $df = 45$, $p < 0.001$, QDS: $\chi^2 = 1,743.5$, $df = 190$, $p < 0.001$). Therefore, the data were appropriate for use in the factor analysis. ODI items were loaded on one factor explaining 51.3 % of the total variance. A four-factor structure (QDS-F1-4) of the Hungarian QDS was found. The four factors explained 71.0 % of the total variance (Table 2). Cronbach- α were 0.89 for the ODI, and 0.92, 0.88, 0.82 and 0.79 for the QDS-F1, -F2, -F3 and -F4, respectively. We also computed Cronbach- α for the whole QDS which was 0.95. No subject scored the possible minimum (0) or maximum (100) score of the ODI or the QDS. When the end-values were narrowed with the MDC values (see below), the ratio of the floor and the ceiling effects were 1.5 % in the ODI and 0.8 % in the QDS [one (QDS) and two (ODI) subjects scored within the MDC at the upper and lower end of the QDS and ODI scores, respectively]. The continuous variables of the study cohort were normally distributed.

Construct validity

Surgically treated patients scored significantly higher on the ODI and the QDS (sub)scales than non-surgically treated subjects (Table 3). Self-reported pain was significantly higher in the surgical group whereas scores of physical, social and environmental subscales of WHOQoL were lower in this treatment group. The ODI and QDS scores, as well as the QDS subscales, were significantly associated with the presence of nerve root deficit (Table 4). Likewise, the ODI and QDS total scores, as well as the QDS subscales, correlated both significantly and highly with pain and the physical subscale of WHOQoL ($p < 0.001$ and $r > 0.4$) (Table 5). The correlation between the ODI and the QDS was excellent ($r = 0.81$, $p < 0.001$) (Fig. 1).

Reproducibility

The mean (SD, range) time between the test and the retest was 9.3 (4.6, 2–14) days. There was no significant difference between the mean ODI and QDS scores when the test and retest measurements were compared (Table 6). Both PROMs showed excellent reproducibility, as confirmed by

Table 2 Internal consistency of the Hungarian QDS (items are sorted according to the factor structure)

Item	Category	Missing answers (%)	Communalities	Factor loadings			
				QDS-F1 ^a	QDS-F2 ^b	QDS-F3 ^c	QDS-F4 ^d
10	Reach up to high shelves	0	0.72	0.70			
11	Throw a ball	6 (4.8)	0.72	0.78			
13	Take food out of the refrigerator	0	0.67	0.79			
14	Make your bed	2 (1.6)	0.72	0.71			
15	Put on socks (panty hose)	0	0.67	0.51			
16	Bend over to clean the bathtub	2 (1.6)	0.67	0.62			
17	Move a chair	0	0.80	0.82			
18	Pull or push heavy doors	0	0.68	0.77			
5	Stand up for 20–30 min	0	0.60		0.72		
7	Climb one flight of stairs	0	0.68		0.57		
8	Walk a few blocks (300–400 m)	1 (0.8)	0.84		0.83		
9	Walk several kilometres	1 (0.8)	0.79		0.83		
12	Run one block (about 100 m)	3 (2.4)	0.67		0.73		
4	Ride in a car	1 (0.8)	0.67			0.58	
6	Sit in a chair for several hours	0	0.73			0.81	
19	Carry two bags of groceries	1 (0.8)	0.68			0.52	
20	Lift and carry a heavy suitcase	2 (1.6)	0.63			0.51	
1	Get out of bed	1 (0.8)	0.70				0.65
2	Sleep through the night	0	0.73				0.83
3	Turn over in bed	2 (1.6)	0.77				0.71
Eigenvalue				9.6	1.8	1.6	1.2
Explained variance (%)				26.2	18.8	13.1	12.9
Cronbach- α				0.92	0.88	0.82	0.79

Factor loadings above 0.5 are shown

^a Everyday activities^b Ambulation^c Sitting/carrying^d Bed/rest**Table 3** Questionnaire scores for surgically and non-surgically treated patients

Variables	Surgical (<i>N</i> = 112) mean (SD)	Non-surgical (<i>N</i> = 21) mean (SD)	<i>p</i> value
ODI	48.4 (16.3)	26.9 (15.7)	0.001
QDS	52.5 (17.1)	30.9 (13.9)	0.001
QDS-F1	16.3 (8.7)	9.3 (6.6)	0.001
QDS-F2	16.6 (5.0)	8.5 (5.7)	0.001
QDS-F3	13.2 (3.8)	9.9 (3.2)	0.001
QDS-F4	6.4 (3.1)	3.1 (2.1)	0.001
Pain	7.1 (1.7)	5.0 (2.5)	0.001
WHOQoL-Phys	55.0 (13.8)	69.2 (14.5)	0.001
WHOQoL-Psy	68.4 (12.8)	69.0 (14.7)	0.831
WHOQoL-Soc	73.3 (14.3)	80.3 (13.3)	0.039
WHOQoL-Env	70.3 (11.1)	77.1 (7.6)	0.008

high ICC values (0.93 and 0.92) and relatively low SEM (4.8 and 5.2 for the ODI and QDS, respectively). The SEM for the two PROMs, expressed in relation to the mean

values of the questionnaires, was about 11 % for both disability questionnaires. MDC_{95%} was approximately 13 points for the ODI and 14 points for the QDS.

Table 4 Questionnaire scores depending on the presence of nerve root deficit

Variables	Nerve root deficit (<i>N</i> = 60) mean (SD)	Without nerve root deficit (<i>N</i> = 73) mean (SD)	<i>p</i> value
ODI	50.6 (15.5)	40.4 (18.6)	0.001
QDS	56.5 (16.3)	43.0 (17.8)	0.001
QDS-F1	18.5 (8.6)	12.5 (7.8)	0.001
QDS-F2	17.3 (4.4)	13.6 (6.5)	0.001
QDS-F3	13.9 (3.3)	11.7 (4.1)	0.002
QDS-F4	6.7 (2.8)	5.1 (3.3)	0.003
Pain	7.4 (1.8)	6.3 (2.1)	0.001
WHOQoL-Phys	52.8 (12.5)	60.9 (15.6)	0.002
WHOQoL-Psy	67.3 (12.9)	69.5 (13.2)	0.338
WHOQoL-Soc	70.4 (14.4)	77.6 (13.6)	0.004
WHOQoL-Env	70.1 (10.8)	72.4 (10.8)	0.233

Table 5 Correlation among pain, WHOQoL subscales and spine related disability

	ODI	QDS	QDS-F1	QDS-F2	QDS-F3	QDS-F4
Pain	0.68*	0.62*	0.49*	0.60*	0.51*	0.57*
WHOQoL-Phys	−0.71*	−0.68*	−0.58*	−0.53*	−0.67*	−0.59*
WHOQoL-Psy	−0.40*	−0.31*	−0.28*	−0.18 [#]	−0.35*	−0.32*
WHOQoL-Soc	−0.31*	−0.26 [#]	−0.28*	−0.19 [#]	−0.21 [#]	−0.23*
WHOQoL-Env	−0.32*	−0.25 [#]	−0.26*	−0.17	−0.22*	−0.23 [#]

Pearson's correlation coefficients (*r*) are represented.
* *p* < 0.001, [#] *p* < 0.01

Fig. 1 Correlation of ODI and QDS scores (*r* 0.81, *p* < 0.001, 95 % CI for the mean)

Reproducibility parameters were also determined for the four QDS subscales (Table 6).

Discussion

The use of patient-reported outcome measurements (PROMs) in spine care and the need for validated, national versions of condition-specific PROMs are widely reported. The goal of this study was to develop validated Hungarian versions of two commonly used back-specific disability

questionnaires, namely the ODI and the QDS. Our results show that the Hungarian versions of these two spine-related PROMs are reliable and valid. Thus, the cross-cultural adaptation of the two questionnaires into Hungarian was successful. During the study, subjects filled in the questionnaires at the hospital in a study office in the presence of a research nurse but the nurse did not help the patient to complete the booklets. In line with ethical requirements, the nurse checked that the personal data form (name, etc.) had been completed before archiving the booklets. She did not give out any further instructions, help the patient to complete the questionnaires or check for missing answers. This process ensured independent completion.

Internal consistency of the Hungarian versions of the ODI and QDS were tested by calculating the Cronbach- α values of the (sub)scales. There is no clear hypothesis with regard to the factor structure of the ODI and the QDS in the literature. This was not published by the original developers and different factor structures have since been published by the few studies examining the factor structure of different national versions of the ODI and QDS. Thus, we performed exploratory factor analyses to determine the factor structure of the Hungarian ODI and QDS. In our study, the ODI items were loaded on one factor. The explained variance by this factor was 51 % and the Cronbach- α of the scale was 0.89, resulting in a positive rating for the internal consistency of the Hungarian ODI [43]. In the literature, we found only two validation studies where

Table 6 Results of test–retest analyses

PROM	Test mean (SD)	Retest mean (SD)	<i>p</i> -value (<i>t</i> -test)	ICC (95 % CI)	SEM	SEM as % of mean ^a	MDC _{95 %}
ODI	45.0 (17.9)	44.6 (17.7)	0.536	0.93 (0.90–0.95)	4.8	10.7	13.3
QDS	49.1 (18.3)	49.8 (18.8)	0.132	0.92 (0.89–0.95)	5.2	10.5	14.4
QDS-F1	15.2 (8.7)	15.8 (8.8)	0.066	0.91 (0.87–0.93)	2.6	16.7	7.2
QDS-F2	15.3 (5.9)	15.4 (5.8)	0.677	0.88 (0.84–0.92)	2.0	13.0	5.5
QDS-F3	12.7 (3.9)	12.6 (4.0)	0.367	0.85 (0.79–0.89)	1.5	11.9	4.2
QDS-F4	5.9 (3.2)	6.0 (3.3)	0.285	0.87 (0.82–0.91)	1.2	20.2	3.3

^a Mean of test and retest

the factor structure of the ODI had been published. Monticone et al. [24] published the one-factor structure of the Italian ODI; the Cronbach- α was 0.86 and the explained variance was 45 %. In contrast, Pekkanen et al. [26] reported that the Finnish version of the ODI had two subscales (factors) explained 51 % of the total variance. Items of the Hungarian QDS loaded on four factors explained 71 % of the total variance in this study. The Cronbach- α values of the four subscales were 0.92 (everyday activities), 0.88 (ambulation), 0.82 (sitting/carrying) and 0.79 (bed/rest), giving a positive rating for internal consistency. The Cronbach- α for the whole scale was 0.95 which was almost the same as that published by Kopeck et al. for the original questionnaire (0.96). The factor structure of the final version of the QDS has been published previously by Christakou et al. [31] only. They found that the items in the Greek version of the QDS loaded on six factors (movement, handling of large/heavy objects, bending/stooping, ambulation, sit/stand and bed/rest). The Cronbach- α values for these factors were between 0.89 and 0.96 which explained 82 % of the total variance. Significant floor and ceiling effects were not observed in the Hungarian versions of the ODI and QDS; not even if the end values of the scales were cut off with the MDC values.

Construct validity of the Hungarian ODI and QDS were analysed by testing pre-defined hypotheses related to the clinical characteristics of our study cohort. In our study, subjects indicated for spinal surgery showed significantly higher disability scores measured by both the ODI and the QDS (sub)scales. The same association was reported by Mannion et al. [21] in a validation study on the German version of the ODI. We also confirmed our other hypothesis that the patients with nerve root deficit (lower extremity sensory disturbance and/or motor weakness) have higher ODI and QDS scores. This is in accordance with the results of Yvanes-Thomas who validated the French version of the QDS [36]. We also analysed correlations between disability, pain and the subscales of WHOQoL-BREF general life quality questionnaire, assuming that the ODI and QDS scores were correlated with pain and the physical subscale of WHOQoL.

Correlations between the ODI and pain have been reported with widely ranging coefficients (r 0.37–0.78) in previously published validation studies that underlined the multi-factorial nature of spine-related disability [44, 45]. We found a very good correlation between the Hungarian ODI and pain (r 0.68) which is similar to that published for the Italian [24], Danish [18], Chinese [19, 20] and German [21] versions. Our correlation between pain and the QDS total score was also very high (r 0.62) and similar to that reported previously for the Dutch [35], Korean [34] and Brazilian-Portuguese [33] versions of the QDS. In addition, we found good correlations amongst the four subscales of the Hungarian QDS and pain (r 0.49–0.6). The Hungarian ODI and QDS scales correlated highly with the physical subscale of the WHOQoL only (r -0.71 and -0.68, respectively, for the ODI and total QDS scores). The same association was reported by Kim et al. [22] who also used this type of general life quality questionnaire in their validation study of the Korean version of the ODI. The four subscales of the Hungarian QDS also correlated only with the physical subscale of the WHOQoL (r 0.53–0.67). The above significant associations with pain, quality of life, neurological signs and indication for treatment provide strong evidence for the construct validity of the Hungarian ODI and QDS.

Reproducibility of the PROMs is often demonstrated by providing a measure of its reliability (i.e. the ICC value of the repeated measurements). Only a minority of the previously reported validation studies of the ODI and the QDS also published the measurement error, expressed as the standard error of measurement (SEM). Using the SEM value, the minimal detectable change (MDC)—the lowest “real” change above the measurement error—can be calculated which is an important feature of the questionnaire [43]. In our study, the MDC_{95%} for the ODI was 13 points which is similar to that reported for the Chinese ODI [20]. It is slightly higher than that reported for the Norwegian (11 points) [29] and German (9 points) [21] versions but lower than the English ODI in comparison studies (15 points) [38, 39]. The MDC for the QDS total score was reported previously by these two comparison studies only. In

comparison to their MDC values (19 [38] and 36 points [39]), our QDS total score MDC_{95%} was lower (14 points); similar to the MDC_{95%} of the Hungarian ODI.

The reliability of the Hungarian ODI and QDS proved to be high in our study. The ICC value for the Hungarian ODI was 0.93. It was higher than that reported for the Finnish [26], Persian [25], Chinese [20], Danish [18], Korean [22], Norwegian [29] and French [17] versions. The ICC for the original (0.99) [12], simplified Chinese (0.99) [19], German (0.96) [21], Italian (0.96) [24], Brazilian-Portuguese (0.99) [28] and Turkish ODI (0.94) [27] was higher than in our study but, in those studies, the mean time between test and retest measurements was lower (1–7 days compared with 9.3 days). Possible memory effect on very high ICC, as published by Fairbank [13], is supported by the fact that the three studies reporting the highest (0.99) ICC values had a 1 day test–retest interval. Consequently, we extended the time gap between the two phases of the study. Similar conclusions can be found in the literature for the ICC of the Hungarian QDS (0.92). Higher ICC was only published by those who used a shorter test–retest period in their studies [33, 37]. The ICC values for the four subscales of the Hungarian QDS were above 0.70 giving a positive rating for the reliability of the four subscales. Considering the high reliability, the MDC_{95%} values above and that fact that the calculation of MDC value also depends on the characteristics of the study cohort, we can conclude that the Hungarian versions of the ODI and QDS reported here can be used as reliable PROMs in future studies on a wide spectrum of spinal patients.

In this study, we performed the translation and validation of two spine-related PROMs in the same study setup. We found an excellent, significant correlation between the Hungarian versions of the ODI and QDS (r 0.81) which was similar to the previously published French (0.9) [17], English (0.82) [39] and (0.80) [14] and Polish (0.82) [23] versions of the questionnaires. Greek [31], Turkish [32] and Korean [34] studies reported weaker correlations between the ODI and the QDS (r : 0.78, 0.67 and 0.72, respectively).

This study has several limitations which should be mentioned. During the linguistic validation process, we followed the instruction of the MAPI research institute. The recommended linguistic validation process of MAPI is different to the widely used method of Beaton et al. [46] in two ways. Firstly, the MAPI methodology does not recommend an informed and an unformed forward translator and two backward translators. In our study, two independent, uninformed translators performed the forward translation of the ODI and the QDS and one independent translator did the backward translation of the PROMs. This deviation in the methodology could limit the use of the Hungarian ODI and QDS despite the attentive further steps of the cross-cultural adaptation process (i.e. expert committee review, pilot testing, clinicians' review). Secondly,

in this study, we did not test the responsiveness of the Hungarian versions of the ODI and QDS. Further studies are needed, therefore, to examine this important psychometric property in Hungarian patients. A further possible limitation is that all the subjects were recruited from the outpatient clinic of the National Center for Spinal Disorders which is a secondary reference centre. Thus, highly disabled patients treated unsuccessfully previously elsewhere could be over-represented in our cohort, as indicated by the relatively high mean scores of the ODI and the QDS and the significant proportion of failed back surgery syndrome patients.

Conclusions

Self-reported patient questionnaires have been reported previously as valid and reliable measurement tools to assess the disability level of spine patients. We have adapted two commonly used PROMs on spine-related disability for use in the Hungarian language. The Hungarian versions of the ODI and QDS proved to be reliable and valid measurement tools that showed good psychometric characteristics. Taking into consideration aforementioned limitations, these adapted questionnaires can be used in everyday clinical practice and scientific studies on Hungarian-speaking spine patients.

Acknowledgments The authors would like to thank Prof. Janec Kopec for useful help in the conceptual definition of the Hungarian QDS and Ms. Zita Bognar for the careful study nursing. The research leading to these results received funding from the European Community's Seventh Framework Programme (FP7, 2007–2013) under grant agreement no. HEALTH-F2-2008-201626 and 269909.

Conflict of interest None.

References

1. Rapp SM (2011) European orthopaedic community steps up its adoption of patient-based outcomes. *Orthopaedics Today Europe* May/June
2. FDA's Guidance on Patient Reported Outcomes in Clinical Research (2009)
3. Costa LO, Maher CG, Latimer J (2007) Self-report outcome measures for low back pain: searching for international cross-cultural adaptations. *Spine (Phila Pa 1976)* 32:1028–1037
4. Fitzpatrick R, Adam S, Ross H et al (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 2(14):1–74
5. Birkenmaier C, Suess O, Pfeiffer M et al (2010) The European multicenter trial on the safety and efficacy of guided oblique lumbar interbody fusion (GO-LIF). *BMC musculoskelet disord* 11:199
6. Zweig T, Mannion AF, Grob D et al (2009) How to Tango: a manual for implementing Spine Tango. *Eur spine j* 18(Suppl 3):312–320

7. Hoy D, Brooks P, Blyth F et al (2010) The Epidemiology of low back pain. *Best pract Res Clin Rheumatol* 24:769–781
8. Katz JN (2006) Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J Bone Joint Surg Am* 88(Suppl2):21–24
9. Pietikäinen S, Silventoinen K, Svedberg P et al (2011) Health-related and sociodemographic risk factors for disability pension due to low back disorders: a 30-year prospective Finnish Twin Cohort Study. *J Occup Env Med* 53:488–496
10. Horváth G, Koroknai G, Acs B et al (2010) Prevalence of low back pain and lumbar spine degenerative disorders. Questionnaire survey and clinical-radiological analysis of a representative Hungarian population. *Int Orthop* 34:1245–1249
11. Muller U, Duetz MS, Roeder C et al (2004) Condition-specific outcome measures for low back pain. Part I: validation. *Eur Spine J* 13:301–313
12. Fairbank JC, Couper J, Davies JB et al (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66:271–273
13. Fairbank JC, Pynsent PB (2000) The Oswestry disability index. *Spine* 25:2940–2952 discussion 2952
14. Kopec JA, Esdaile JM, Abrahamowicz M et al (1995) The Quebec back pain disability scale. Measurement properties. *Spine* 20:341–352
15. Kopec JA, Esdaile JM, Abrahamowicz M et al (1996) The Quebec back pain disability scale: conceptualization and development. *J Clin Epidemiol* 49:151–161
16. Payares K, Lugo LH, Morales V et al (2011) Validation in Colombia of the Oswestry disability questionnaire in patients with low back pain. *Spine* 36:E1730–E1735
17. Denis I, Fortin L (2012) Development of a French-Canadian Version of the Oswestry Disability Index. *Spine* 37:E439–E444
18. Lauridsen HH, Hartvigsen J, Manniche C et al (2006) Danish version of the Oswestry Disability Index for patients with low back pain. Part 1: cross-cultural adaptation, reliability and validity in two different populations. *Eur Spine J* 15:1705–1716
19. Liu H, Tao H, Luo Z (2009) Validation of the simplified Chinese version of the Oswestry disability index. *Spine* 34:1211–1216 discussion 1217
20. Lue Y-J, Hsieh C-L, Huang M-H et al (2008) Development of a Chinese version of the Oswestry disability index version 2.1. *Spine* 33:2354–2360
21. Mannion AF, Junge A, Fairbank JCT et al (2006) Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability, and validity. *Eur Spine J* 15:55–65
22. Kim D-Y, Lee S-H, Lee H-Y et al (2005) Validation of the Korean version of the Oswestry disability index. *Spine* 30:E123–E127
23. Misterska E, Jankowski R, Glowacki M (2011) Quebec back pain disability scale, low back outcome score and revised Oswestry low back pain disability scale for patients with low back pain due to degenerative disc disease: evaluation of Polish versions. *Spine* 36(26):E1722–E1729
24. Monticone M, Baiardi P, Ferrari S et al (2009) Development of the Italian version of the Oswestry Disability Index (ODI-I): a cross-cultural adaptation, reliability, and validity study. *Spine* 34:2090–2095
25. Mousavi SJ, Parnianpour M, Mehdian H et al (2006) The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec back pain disability scale: translation and validation studies of the Iranian versions. *Spine* 31:E454–E459
26. Pekkanen L, Kautiainen H, Ylinen J et al (2011) Reliability and validity study of the Finnish version 2.0 of the Oswestry disability index. *Spine* 36:332–338
27. Yakut E, Düger T, Oksüz C et al (2004) Validation of the Turkish version of the Oswestry Disability Index for patients with low back pain. *Spine* 29:581–585 discussion 585
28. Vigatto R, Alexandre NMC, Correa Filho HR (2007) Development of a Brazilian Portuguese version of the Oswestry Disability Index: cross-cultural adaptation, reliability, and validity. *Spine* 32:481–486
29. Grotle M, Brox JI, Vøllestad NK (2003) Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *J Rehabil Med* 35:241–247
30. Boscainos PJ, Sappas G, Stilianessi E et al (2003) Greek versions of the Oswestry and Roland-Morris Disability Questionnaires. *Clin Orthop Relat Res* 411:40–53
31. Christakou A, Andriopoulou M, Asimakopoulos P (2011) Validity and reliability of the Greek version of the Quebec back pain disability scale. *J Back Musculoskeletal Rehabil* 24:145–154
32. Melikoglu MA, Kocabas H, Sezer I et al (2009) Validation of the Turkish version of the Quebec back pain disability scale for patients with low back pain. *Spine* 34:E219–E224
33. Rodrigues MF, Michel-Crosato E, Cardoso JR et al (2009) Psychometric properties and cross-cultural adaptation of the Brazilian Quebec back pain disability scale questionnaire. *Spine* 34:E459–E464
34. Suh KT, Kim JI, Lim JM et al (2011) Validation of the Korean version of the Quebec back pain disability scale. *J Spinal Disord Tech* 25(8):447–450
35. Schoppink LE, van Tulder MW, Koes BW et al (1996) Reliability and validity of the Dutch adaptation of the Quebec back pain disability scale. *Phys Ther* 76:268–275
36. Yvanes-Thomas M, Calmels P, Béthoux F et al (2002) Validity of the French-language version of the Quebec back pain disability scale in low back pain patients in France. *Joint Bone Spine* 69:397–405
37. Bendeddouch I, Rostom S, Bahiri R et al (2012) Translation, adaptation and validation of the Moroccan version of the Quebec back pain disability scale. *Clin Rheumatol* 31(6):943–949
38. Davidson M, Keating JL (2002) A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 82:8–24
39. Fritz JM, Irrgang JJ (2001) A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec back pain disability scale. *Phys Ther* 81:776–788
40. Muller U, Roeder C, Dubs L et al (2004) Condition-specific outcome measures for low back pain. Part II: scale construction. *Eur Spine J* 13:314–324
41. Paulik E, Belec B, Molnar R et al (2007) Applicability of the abbreviated version of the World Health Organization's quality of life questionnaire in Hungary. *Orv Hetil* 148:155–160
42. Methodology for linguistic validation (MAPI guideline). <http://www.mapi-institute.com/linguistic-validation/methodology>
43. Terwee CB, Bot SDM, de Boer MR et al (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42
44. Mannion AF, Junge A, Taimela S et al (2001) Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy. *Spine (Phila Pa 1976)* 26:920–929
45. Grotle M, Vøllestad NK, Veierød MB et al (2004) Fear-avoidance beliefs and distress in relation to disability in acute and chronic low back pain. *Pain* 112:343–352
46. Beaton DE, Bombardier C, Guillemin F, Ferraz MB (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 25(3186):3191