

H. Koçyigit · Ş. Gülseren · A. Erol · N. Hizli
A. Memis

The reliability and validity of the Turkish version of Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO)

Received: 17 May 2002 / Accepted: 31 July 2002
© Clinical Rheumatology 2003

Abstract The purpose of this study was to investigate the reliability and validity of the Turkish version of Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO). The patient group included 43 females aged between 55 and 78 years with vertebral fractures due to osteoporosis. The control group consisted of 43 healthy female volunteers whose ages matched those of the patients. All of the participants were evaluated using both QUALEFFO and SF-36. In the reliability studies, internal consistency within the domain of QUALEFFO was generally good, with Cronbach's α values ranging between 0.70 and 0.96. Convergent and discriminant validity rates of domains were both found to be between 89% and 100%. Significant correlations existed between scores of similar domains of QUALEFFO and the SF-36, especially for pain, physical function, social function and general health perception. The receiver operating characteristic (ROC) curve analysis of QUALEFFO and the SF-36 indicated that all five domains in each questionnaire were significantly predictive of vertebral fractures. However, when comparing similar domains of the two questionnaires, the social function domain of QUALEFFO demonstrated a significantly better performance. In conclusion, the Turkish version of QUALEFFO was found to be reliable and valid in the evaluation of patients with vertebral fractures due to osteoporosis. Our study also suggests that the patients with vertebral fractures due to osteoporosis have impairment in quality of life.

Keywords Osteoporosis · Quality of life · Questionnaires · Spinal fractures

Introduction

In a world where the average human life span keeps on extending, chronic diseases of the elderly are a fundamental issue. One of these diseases is osteoporosis, and it is gradually drawing more attention in many countries, including Turkey. Unfortunately, something that is overlooked in the treatment of osteoporosis is the concept of health-related quality of life [1,2]. It is inevitable that today physicians pay more attention to measures of quality of life that imply patients' ideas and emotions about their illness, along with important investigations such as physical examination and imaging techniques. Although medical measurements provide important information to physicians they mean less to patients. The results of these methods seldom correlate with the concepts of 'wellbeing' and 'functional capacity' to which patients pay greater attention.

The clinical importance of osteoporosis is determined by subsequent fractures that lead to impairment in quality of life [1,2]. Vertebral fractures due to osteoporosis are not often symptomatic and may pass unnoticed by physicians. It is estimated that only one-third of patients present to physicians or hospitals when fracture occurs [3]. Vertebral fractures occur 10 times more often in women than in men [4]. The prevalence of vertebral osteoporosis has been estimated to be between 40% and 45% in Turkey [5]. Vertebral fractures may cause acute and chronic back pain. They may also lead to immobility and impairment in daily activities and mental function [6]. Recently several instruments for measuring the quality of life of patients with osteoporosis have been developed. One of them is the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO), which has been developed by a working party of the European Foundation for Osteoporosis (EFO) for patients with vertebral fractures [7,8]. This questionnaire has been translated into seven different languages and has been found to be valid [8].

H. Koçyigit (✉) · Ş. Gülseren · A. Erol · N. Hizli · A. Memis
Izmir Atatürk Teaching Hospital, Izmir, Turkey
1408 Sokak No: 2/4, Alsancak, 35220 Izmir, Turkey
Tel.: +90.533 637 3839
Fax: +90 232 463 5273
E-mail: drhikmet@hotmail.com

To be used internationally, these kinds of questionnaire must be translated into the respective local languages and must also be culturally adapted. The measurement properties of the instrument must then be re-evaluated to ensure that it retains the original version's properties. The purpose of this study was to evaluate the reliability and validity of the Turkish version of QUALEFFO.

Patients and methods

The patient group consisted of 43 women with clinical osteoporosis who had vertebral fractures. Inclusion criteria were age between 55 and 78 years, lumbar bone mineral density (BMD) T-score < -1 and at least one vertebral fracture (reduction of anterior, middle or posterior vertebral height of more than 20% on clinical reading). In addition, the patients had to be able to complete the questionnaire. Patients who had any important physical disorder were not included in the study. Patients with recent vertebral or other fractures (within 1 month), patients with metabolic bone disease and those with malignancies were excluded. Patients who had in recent years received treatment with medications known to influence bone metabolism (such as calcitonin, alendronate etc.) were also excluded.

Control subjects were 43 healthy women chosen from the general population and were age-matched to patients. Subjects with chronic back pain and vertebral or other fractures were excluded. Subjects with conditions exerting a major influence on quality of life, such as malignancies, were also excluded.

Questionnaires

QUALEFFO is a self-assessment instrument that has been developed to measure the quality of life of patients with vertebral osteoporosis. It consists of 41 questions and five domains. These are: pain, physical function, social function, general health perception and mental function. It takes about 10 minutes for the patient to complete.

First, the English version of QUALEFFO was translated into Turkish by a physiatrist and a psychiatrist whose native language was Turkish and who spoke English. Later, this version was independently translated back into English by both a native English-speaking professor and a rheumatologist who spoke English. All the texts were compared with each other and the final Turkish version was documented. It was reprinted by kind permission of Dr P. Lips; Working Party of EFFO, Academic Hospital Vrije University, Amsterdam, The Netherlands.

Along with QUALEFFO, the Short Form 36 of the Medical Outcome Study (SF-36) was also administered to test validity. SF-36 is a generic instrument developed for measuring quality of life. It is not specific for any disease. Like QUALEFFO, SF-36 is a self-assessment instrument and it can be completed in about 5 minutes. It consists of eight domains: bodily pain, physical functioning, social functioning, general health, mental health, vitality, and role restrictions due to physical and emotional problems. The adaptation studies of SF-36 for use in Turkey were made by Kocyigit et al [9].

Radiographs

Standardised lateral radiographs of the dorsal and lumbar spine were taken centred on T8 and L3, with a film focal distance of 105 cm. Vertebral deformity was defined when anterior, middle, posterior or overall height loss was more than 20%. In addition, the number of deformities was recorded for each patient.

Conduct of the study

All the subjects included in the study answered the questionnaires in the same order: QUALEFFO followed by SF-36. The

questionnaires were again administered after 2 weeks in order to assess test-retest reliability.

Statistical analysis

The answers for each question in QUALEFFO were scored from 1 to 5, except for Questions 23, 24, 25 and 26 (scored 1–3), and Questions 27, 28 and 29 (scored 1–4). For Questions 24, 26 and 29, the answers 'not applicable', and 'no cinema or theatre within a reasonable distance' were not scored. The scores for Questions 33, 34, 35, 37, 39 and 40 were reversed so that a low score always indicates better health and a high score always indicates worse health. Domain scores were calculated by summing the scores of questions included in the domain and submitting the sum to a linear transformation to a scale of 100. Zero indicated good health status and 100 indicated poor health status. A total QUALEFFO score was computed by summing the scores of all questions and submitting the sum to a linear transformation to a scale of 100. SF-36 was scored according to the instructions in the guide [10]. SF-36 was also scored from 0 to 100 but, in contrast, 0 indicated poor health status and 100 indicated good health status.

Statistical Package for the Social Sciences (SPSS v. 10.00) was used for the statistical analysis. In reliability studies for the Turkish version of QUALEFFO, Cronbach's α coefficient was used to assess internal consistency. This was calculated not only for the entire QUALEFFO but also for each domain. To assess test-retest reliability, a κ score was determined for each question. To assess validity, convergent validity and discriminant validity were studied. When the correlation coefficient between the score for each question and its own total domain score was higher than 0.40, it was agreed that there was adequate convergent validity. When the correlation coefficient of the score for each question with its own domain score was higher than with total scores of other domains, it was agreed that there was adequate discriminant validity. In addition, the correlation between scores of the similar domains of QUALEFFO and SF-36 was studied. As the mental function domain of QUALEFFO also contains questions about vitality, mental health and vitality domains, the scores of the SF-36 were united during comparison. Neither of the two role functioning domains of the SF-36 was included in this between-domain comparison because they did not have corresponding QUALEFFO domains. Receiver operating characteristic (ROC) curves were constructed (MedCalc package version 6.16) to compare the ability of QUALEFFO and SF-36 domains to discriminate between cases and controls over all possible cutoff values of the questionnaire scores [11]. The significance of differences between areas under curves was used to compare domains of QUALEFFO and SF-36. In all statistical analyses, a value of correlation coefficient between 0 and 0.25 was regarded as 'no or very poor' correlation; 0.25–0.50 was regarded as 'poor-moderate' correlation; 0.50–0.75 was regarded as 'good' correlation and 0.75–1.00 was regarded as 'very good' correlation.

Results

Forty-three patients who had been followed up in Izmir Atatürk Teaching Hospital, Department of Physical Therapy and Rehabilitation, Division of Osteoporosis, and 43 age-matched control subjects chosen from the general population, were studied. The mean ages of patients and control subjects were 65.3 years (SD 5.7 years) and 64.1 years (SD 5.5 years), respectively. All subjects were female and Caucasian. There was no significant difference between the two groups with respect to the duration of education. The mean number of years of education of patients and control subjects was 8.3

years (SD 1.9 years) and 8.1 years (SD 2.2 years), respectively. According to the mentioned criteria, mean number of vertebral fractures was 2.28 (SD 1.22) in patients; 14 patients (33%) had one fracture, 12 patients (28%) had two fractures and 17 (39%) had three or more fractures. The fractures were mostly detected in vertebral regions D10, D11 and L1. None of the control subjects had fractures.

In reliability studies for the Turkish version of QUALEFFO, Cronbach's α was calculated for each domain and values between 0.70 and 0.96 were obtained (Table 1). Cronbach's α was 0.97 for the entire QUALEFFO, and this high value was indicative of a good level of internal consistency. When test-retest reliability was studied, κ scores ranging from 0.46 to 0.87 were obtained for 41 questions, and for 23 of them κ was greater than 0.70.

In validity studies, item-domain total score correlation coefficients were calculated for each domain, with each item the domain contains. Item-domain total score correlation coefficients were between 0.775 and 0.902 for pain, 0.545 and 0.888 for physical function, 0.443 and 0.850 for social function, 0.825 and 0.852 for general health perception, and between 0.269 and 0.751 for mental function. Item-unassociated domain total score correlations were generally lower than item-associated domain total score correlations. Convergent and discriminant validity ratios are shown in Table 2.

Mean scores on QUALEFFO and SF-36 in patients with vertebral fractures and control subjects are presented in Table 3. Compared to those of control subjects the scores in patients with vertebral fractures due to osteoporosis clearly show significant impairment in quality of life.

Table 1 Cronbach's α coefficients of each domain of the Turkish version of the QUALEFFO

Domain (no of question)	Internal consistency (Cronbach's α)
Pain (5)	0.90
Physical function (17)	0.96
Social function (7)	0.80
General health perception (3)	0.79
Mental function (9)	0.70
For entire QUALEFFO	0.97

Table 2 Convergent and discriminant validity ratios of each domain of the Turkish version of the QUALEFFO

Domain	Convergent validity (%)	Discriminant validity (%)
Pain	100	100
Physical function	100	100
Social function	100	100
General health perception	100	100
Mental function	89	89

When correlations between corresponding domains of QUALEFFO and SF-36 were studied Spearman correlation coefficients were significantly high (Table 4). When patients were grouped according to the number of fractures, no statistically significant correlation was found between the QUALEFFO domain scores and the number of fractures.

Figure 1 demonstrates the ROC curve analysis for the QUALEFFO domains and the QUALEFFO total score, which was carried out to evaluate the discriminatory capacity between fracture cases and controls. All of the QUALEFFO domains and the QUALEFFO total score were significantly predictive of vertebral fracture. The ROC curve analyses for QUALEFFO, SF-36, and the differences between the two questionnaires are summarised in Table 5. The three most discriminatory domains were pain, physical function and general health perception. However, all of the domains within each questionnaire were significantly predictive of vertebral deformity. The discriminant power of the total QUALEFFO score (area = 0.93, SE = 0.03; $P < 0.001$) was also good. When each domain of QUALEFFO was compared with the corresponding domain of the SF-36, only the social function domain of QUALEFFO demonstrated significantly superior performance ($P = 0.001$). The results for the other domains did not differ significantly between the two questionnaires.

Discussion

QUALEFFO has been especially developed to consider the problems of patients with vertebral fractures due to osteoporosis could have during daily activities. Generally it can be said that these kinds of questionnaires are new for our culture. Despite this, neither the patients nor the controls had difficulty completing the questionnaire. However, as this questionnaire is easy to understand and appropriate for our culture, we suppose that subjects can complete the questionnaire in a short time without difficulty. High κ scores obtained from test-retest reliability studies support this observation.

In functional assessment scales, Cronbach's α values above 0.70 are agreed to be adequate for internal consistency [12]. In our study, internal consistency measured by Cronbach's α for each domain ranged from 0.70 to 0.96. Lips et al. [8], who developed the questionnaire, have found Cronbach's α coefficients of domains between 0.72 and 0.92. In the study of Murrell et al. [13] the internal consistency of QUALEFFO was assessed and Cronbach's α values of the domains ranged from 0.70 to 0.91. In the study of Randell et al. [14] the internal consistency of another osteoporosis targeted questionnaire, the Osteoporosis Assessment Questionnaire (OPAQ), was assessed and Cronbach's α values of the scales ranged from 0.72 and 0.92. In our study, the coefficient of internal consistency was 0.70 for mental function domain and this value was lower than

Table 3 Scores on the QUALEFFO and the SF-36 domains in patients and controls

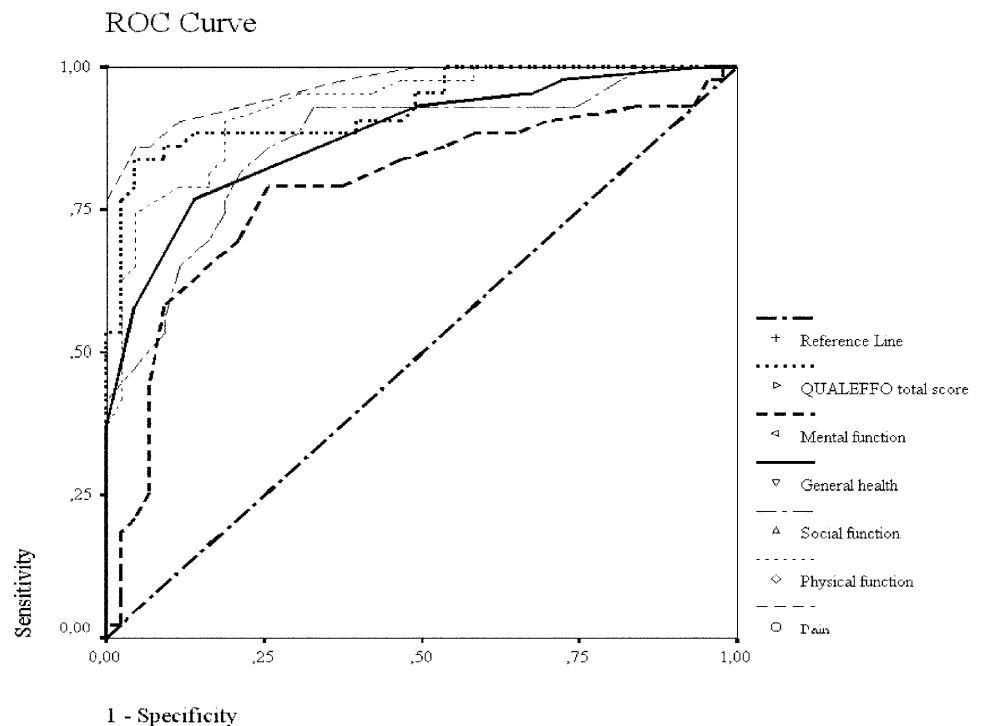
Domain		Controls	Patients	<i>t</i> value	<i>P</i> value
		Mean ± SD	Mean ± SD		
Q	Pain	21.6 ± 14.4	63.4 ± 16.1	12.65	0.001
U	Physical function	22.8 ± 13.5	54.2 ± 16.5	9.66	0.001
A	Social function	43.3 ± 17.6	67.8 ± 14.6	7.01	0.001
L	General health per.	50.2 ± 13.4	71.5 ± 12.5	7.62	0.001
E	Mental function	43.7 ± 12.7	56.9 ± 14.5	4.48	0.001
F					
F	Total QUALEFFO score	31.9 ± 11.7	58.9 ± 12.5	10.32	0.001
O					
S	Bodily pain	73.2 ± 16.0	32.9 ± 13.1	12.76	0.001
F	Physical function	66.7 ± 11.3	39.2 ± 17.7	8.59	0.001
-	Social function	56.4 ± 20.5	40.7 ± 16.6	3.90	0.001
3	General health	57.4 ± 14.2	35.8 ± 15.1	6.82	0.001
6	Mental health and vitality	57.8 ± 11.1	48.8 ± 15.4	3.08	0.003

Table 4 Spearman's rank correlation coefficients between scores of similar domains of QUALEFFO and SF-36 instruments

QUALEFFO domain	SF-36 domain	Correlation coefficient*	<i>P</i> value
Pain	Bodily pain	-0.93	0.001
Physical function	Physical functioning	-0.86	0.001
Social function	Social functioning	-0.75	0.001
General health perception	General health	-0.81	0.001
Mental function	Mental health and vitality	-0.68	0.001

* As the two scales run in the opposite direction, negative values have been obtained

Fig. 1 Receiver operating characteristic (ROC) curve for total QUALEFFO score and individual QUALEFFO domain performance in discriminating between vertebral fracture cases and non-fractured controls



the values calculated for other domains. Leave-one-out diagnostics indicated that Question 39 and Question 41 were most responsible for this relatively low Cronbach's α . These questions are, respectively, 'Do you find it easy

to make contact with people?' and 'Are you afraid of becoming totally dependent?' This situation may be due to difficulties in understanding and answering these two questions because of cultural factors. Nevertheless, both

Table 5 Receiver operating characteristic (ROC) curve analysis for QUALEFFO and the SF-36

Domain	QUALEFFO area* (SE)	SF-36 area* (SE)	Difference area (SE)	<i>P</i> value
Pain	0.97 (0.02)	0.98 (0.02)	-0.01 (0.02)	0.332
Physical function	0.93 (0.03)	0.90 (0.04)	0.03 (0.03)	0.222
Social function	0.87 (0.04)	0.73 (0.06)	0.14 (0.04)	0.001
General health perception	0.88 (0.04)	0.85 (0.04)	0.03 (0.03)	0.406
Mental function	0.79 (0.05)	0.72 (0.06)	0.07 (0.05)	0.196

The null hypothesis states that the questionnaire has no discriminant power (area under the curve = 0.5, line of unity). Higher values indicate better discrimination. The *P* value refers to the difference between the two questionnaires

* All values are significantly different from 0.5 ($P < 0.001$)

two items had higher correlation with mental function domain than with the other domains. As a result, both items should be evaluated as part of mental function domain. On the other hand, also for face validity, these two items should better be evaluated in mental function domain. However, these results indicate that internal consistency of the Turkish version of QUALEFFO is both adequate and high.

In validity studies the convergent validity ratios of domains were between 89% and 100%, discriminant validity ratios were also between 89% and 100%. Lips et al. reported the same ratios as 78%–100% for convergent validity and 72%–100% for discriminant validity [8]. Our findings also indicate that the validity of the questionnaire is adequate.

This study suggests that the quality of life in patients with vertebral osteoporosis is significantly impaired compared with that of age-matched control subjects. As shown in Table 3, when mean scores of corresponding domains of generic instrument SF-36 were taken into consideration, it was seen that patient and control groups differed significantly from each other. When mean scores of domains of specific instrument QUALEFFO were taken into consideration, the two groups differed significantly again. On the other hand, corresponding domains of the two instruments were highly correlated with each other in our study. These findings suggest that in discriminating between patients and controls, QUALEFFO is similar in structure to an established general health status measure, the SF-36. In this respect, our findings were quite consistent with those of Murrell et al. [13]. In their study, they observed the good psychometric properties of the QUALEFFO in terms of excellent test-retest reliability, good internal consistency and very acceptable levels of validity when related to the SF-36.

According to the results of ROC analyses, all of the domains of both QUALEFFO and the SF-36 were significantly predictive of vertebral deformity. However, the most discriminatory domains of both instruments were pain, physical function and general health perception. The findings in our study were equivalent to those reported by Lips et al. [8]. Consequently, it can be said

that both instruments performed very well in discriminating vertebral fracture cases from non-fracture controls. In this study, there was no significant correlation between the scores of QUALEFFO domains and the number of vertebral fractures. Lips et al. [8] and Cantarelli et al. [15] also found no correlation between the number of fractures and the components of QUALEFFO or the OPAQ.

The results of this study show that the Turkish version of QUALEFFO is repeatable, reliable and valid in the evaluation of patients with vertebral fractures due to osteoporosis that have high treatment costs. In addition, our findings confirm that quality of life is impaired in patients with vertebral fractures due to osteoporosis.

This paper was presented as a poster at the 23rd Annual Meeting of American Society for Bone and Mineral Research, Phoenix, Arizona, USA, October 12–16, 2001.

References

1. Gold DT (2001) The nonskeletal consequences of osteoporotic fractures. Psychologic and social outcomes. *Rheum Dis Clin North Am* 27: 255–62
2. Oleksik A, Lips P, Dawson A et al. (2000) Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *J Bone Miner Res* 15: 1384–92
3. O'Neill TW, Felsenberg D, Varlow J et al. (1996) The prevalence of vertebral deformity in European men and women: The European Vertebral Osteoporosis Study. *J Bone Miner Res* 11: 1010–8
4. Lindsay R (1988) Pathogenesis, detection and prevention of postmenopausal osteoporosis. In: Studd J, Whitehead M, eds. *The menopause*. Oxford: Blackwell Scientific Publications 156–67
5. Eryavuz M (1998) Osteoporoz epidemiyolojisi. In: Kutsal YG, ed. *Osteoporoz*. Istanbul, 8–32
6. Kanis JA, Minne HW, Meunier PJ et al. (1992) Quality of life and vertebral osteoporosis. *Osteoporosis Int* 2: 161–3
7. Lips P, Cooper C, Agnusdei D et al. (1997) Quality of life as outcome in the treatment of osteoporosis: the development of a questionnaire for quality of life by the European Foundation for Osteoporosis. *Osteoporosis Int* 7: 36–8
8. Lips P, Cooper C, Agnusdei D et al. (1999) Quality of life in patients with vertebral fractures: Validation of the quality of life questionnaire of the European Foundation for Osteoporosis (QUALEFFO). *Osteoporosis Int* 10: 150–60

9. Koçyiğit H, Aydemir Ö, Fisek G et al. (1999) Kısa Form-36 (KF-36)'nin Türkçe versiyonunun güvenilirliği ve geçerliliği. *İlaç ve Tedavi Dergisi* 12: 102–6
10. Medical Outcomes Trust (1993) How to score the SF-36 Health Survey. Boston, MA: The Medical Outcomes Trust, 1–25
11. www.medcalc.be
12. Granger CV, Cotter AC, Hamilton BB et al. (1990) Functional assessment scales: a study of persons with multiple sclerosis. *Arch Physical Med Rehab* 71: 870–5
13. Murrell P, Todd CJ, Martin A et al. (2001) Postal administration compared with nurse-supported administration of the QUALEFFO-41 in a population sample: Comparison of results and assessment of psychometric properties. *Osteoporosis Int* 12: 672–9
14. Randell AG, Bhalerao N, Nguyen TV et al. (1998) Quality of life in osteoporosis: reliability, consistency and validity of the Osteoporosis Assessment Questionnaire. *J Rheumatol* 25: 1171–9
15. Cantarelli FB, Szejnfeld VL, Oliveira LM et al. (1999) Quality of life in patients with osteoporosis fractures: Cultural adaptation, reliability, and validity of the Osteoporosis Assessment Questionnaire. *Clin Exp Rheumatol* 17: 547–51