



<b>Title</b>	<b>Validation of the quality of life questionnaire of the European foundation for osteoporosis (QUALEFFO-31) in Chinese</b>
<b>Author(s)</b>	<b>Lai, BMH; Tsang, SWY; Lam, CLK; Kung, AWC</b>
<b>Citation</b>	<b>Clinical Rheumatology, 2010, v. 29 n. 9, p. 965-972</b>
<b>Issued Date</b>	<b>2010</b>
<b>URL</b>	<b><a href="http://hdl.handle.net/10722/144965">http://hdl.handle.net/10722/144965</a></b>
<b>Rights</b>	<b>The Author(s)</b>

# Validation of the quality of life questionnaire of the European foundation for osteoporosis (QUALEFFO-31) in Chinese

Billy M. H. Lai · Shirley W. Y. Tsang ·  
Cindy L. K. Lam · Annie W. C. Kung

Received: 4 November 2009 / Revised: 30 April 2010 / Accepted: 16 May 2010 / Published online: 26 June 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** QUALEFFO-31 is a recently developed disease-specific instrument derived from QUALEFFO-41 and intended to have improved efficacy and response rates. We aimed to validate QUALEFFO-31 in Chinese and examine the use of QUALEFFO-31 in clinical practice. This questionnaire was translated into Chinese and applied to 118 case–control pairs aged between 50 and 85 years with prevalent osteoporotic vertebral fractures to evaluate its validity, repeatability, and discriminatory ability. It was also used to evaluate the quality of life (QOL) of 69 case–control trios with prevalent clinical and morphometric fractures. The QOL of all subjects was concurrently assessed using SF-36 for comparison. QUALEFFO-31 had good internal consistency with adequate convergent and discriminatory validity. The median test–retest repeatability ranged from 0.65–0.85. In general, there were good correlations between QUALEFFO-31 and SF-36. ROC curve analysis revealed that QUALEFFO-31 had significant ability to discriminate between clinical fracture subjects versus morphometric fracture subjects and controls. QUALEFFO-31 also demonstrated higher discriminatory capacity for pain. Subjects with clinical vertebral fractures (CVFs) had a significant reduction in QOL compared with

other subjects. The QUALEFFO-31 is a useful tool for assessing QOL in Chinese. It was well accepted and significantly predictive of subjects with CVFs.

**Keywords** Chinese · Construct validation · Health-related quality of life · Morphometric fracture · Osteoporosis · QUALEFFO-31

## Introduction

For the past few decades, osteoporosis has become a growing public health concern worldwide associated with aging of the world's population [1]. It has been estimated that 30–50% of women and 15–30% of men will suffer an osteoporotic fracture in their lifetime [2]. The most common clinical complications of osteoporosis are hip fracture, vertebral deformity, and wrist fracture. Although hip fractures are less prevalent in Asia than the USA [3, 4], vertebral fractures are as frequent in Asian as in Caucasian women [5]. Among Southern Chinese in Hong Kong, the prevalence of vertebral fracture has been reported to be 17% in men and 30% in women [3, 6]. Vertebral fractures can be classified as clinical (symptomatic) or morphometric (radiographic) fractures. Both may be associated with significant morbidity in terms of physical and psychosocial functioning and reduce a subject's health-related quality of life (HRQoL) [7–12]. The concept of HRQoL refers to a person or group's perceived physical and mental health over time. In recent years, a growing number of health care professionals use HRQoL to estimate the burden of disease in a population and to compare the consequences of different diseases. Instruments that measure HRQoL are thus an essential means by which to determine intervention strategies and treatment outcomes in these patients.

---

Billy M. H. Lai and Shirley W. Y. Tsang contributed equally.

B. M. H. Lai · S. W. Y. Tsang · A. W. C. Kung (✉)  
Department of Medicine, Queen Mary Hospital,  
The University of Hong Kong,  
Pokfulam Road,  
Hong Kong, China  
e-mail: awckung@hku.hk

C. L. K. Lam  
Family Medicine Unit, Ap Lei Chau Clinic,  
The University of Hong Kong,  
Hong Kong, China

Generic or disease-specific instruments can be used to assess HRQoL: traditionally, clinicians have relied on generic instruments (such as the Short-Form 36, Sickness Impact Profile, and the Nottingham Health Profile) in subjects with vertebral fractures although such instruments fail to explore the specific aspects of osteoporosis. Disease-specific instruments for osteoporosis have therefore recently gained favor because they have greater face and content validity for the individual disease. The Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO-41) is a well-established disease-specific tool for assessing HRQoL in subjects with clinical [9] and morphometric [13] vertebral fractures. At present, this questionnaire has been translated into 20 languages and has been found to be validated for use in other countries with diverse cultures including Turkey [14] and Mexico [15]. QUALEFFO-41 consists of 41 questions in five domains (pain, physical functioning, social functioning, general health perception, and mental functioning). Nonetheless the large number of items included in QUALEFFO-41 limited its clinical application so QUALEFFO-31 was developed. Derived from QUALEFFO-41, the shorter questionnaire that comprised QUALEFFO-31 was intended to reap better response rates and to be more efficacious in clinical practice. Although QUALEFFO-31 [16] was derived by factor analysis on a sound theoretical basis based on studies conducted in Western populations, this new questionnaire has yet to be validated and implemented in clinical practice in Asian populations. To allow comparisons across countries, we translated QUALEFFO-31 into Chinese and applied it to a local population in Hong Kong in order to validate the cross-cultural adaptation of the questionnaire for use in Chinese populations. Therefore, the aims of this study were [1] to assess the reliability and validity of QUALEFFO-31 in Chinese populations in Hong Kong and [2] to investigate the capacity of QUALEFFO-31C and SF-36 to differentiate between subjects with clinical vertebral fractures and those with morphometric fracture.

## Methods

### Subjects

Three groups of subjects were recruited: clinical fracture cases, morphometric fracture cases, and controls. Subjects presented clinically with low back pain to Queen Mary Hospital and identified to have vertebral fractures on spine X-ray were recruited as clinical fracture cases. Additional inclusion criteria included age between 50 and 85 years, presence of at least one vertebral fracture, i.e., >20% reduction in vertebral height (anterior, middle, or posterior) on spine X-ray, and possessing lumbar L1–4 spine bone

mineral density T-score  $<-1$ . Subjects had to be ambulant and capable of completing the questionnaire. Subjects were recruited into the study at least 6 months after their fracture when they were considered to be stable. Patients with fractures at other non-spine sites that interfered with pain or activity and those with medical conditions that could affect their activity of daily living and pain were excluded. Each clinical fracture case was matched for sex and age ( $\pm 3$  years) with one control subject who responded to advertisement posted in the hospital. Control subjects with medical conditions that exerted a major influence on the quality of life were excluded from the study. Control subjects were required to have no chronic low back pain, apparent kyphosis, and no morphometric vertebral fractures as defined by the previously stated criteria.

“Morphometric fracture cases” were defined as individuals who displayed radiographic evidence of a vertebral fracture according to Genant’s classification [17] in the absence of clinical osteoporosis as defined above. These were subjects referred to the Osteoporosis Centre for bone mineral density assessment. Other inclusion and exclusion criteria for this group of subjects were the same as before. For comparison of HRQoL, each morphometric fracture case was matched for sex and age ( $\pm 3$  years) with a clinical fracture subject and two control subjects.

Informed consent was obtained from all study subjects and demographic information collected at the baseline visit. Approval was obtained from the Ethics Committee of The University of Hong Kong and all procedures were conducted in line with the terms of the Declaration of Helsinki.

### Questionnaires

The Chinese version of QUALEFFO-31 (QUALEFFO-31C) was developed from the English version and translated into Chinese. It was then translated back to English by blinded individuals using standard procedures to confirm accuracy of translation. The QUALEFFO-31C consists of three domains: [1] pain, [2] physical function, and [3] mental function. Response options are in the form of a five-point ordinal scale, with lower scores indicating better quality of life. The QUALEFFO provides both domain scores and an overall score [9].

The Chinese version of the Short-Form 36 of the Medical Outcome Study (SF-36), validated previously by Lam et al. [18], was used as a reference for comparison with QUALEFFO-31C questionnaires. The Chinese version of the Short-Form 36 covers eight multi-item scales: physical functioning, limitations due to physical health problems, bodily pain, general health, vitality, social functioning, limitations due to emotional health problems, mental health, and one single-item scale on health transition.

Questionnaires were administered to subjects (both cases and controls) in Cantonese, the language for Southern Chinese in Hong Kong, at the Osteoporosis Clinic of Queen Mary Hospital before any investigative procedures were performed. Both questionnaires were completed by interview of the subject. The questionnaires were subsequently administered again to the same subjects after 4 weeks. All inter-current events between the two visits were noted.

#### Radiographic assessment

Standardized lateral radiographs of the thoracic and lumbar vertebrae (T4–L5) were taken and assessed by experienced radiologists to identify vertebral deformities according to Genant's classification [17].

#### Statistical analysis

Statistical analysis was based on the protocol used by Lips et al. in the validation study of QUALEFFO-41 [9]. The scoring algorithm provided by the International Osteoporosis Foundation for the QUALEFFO was used. A lower score indicates better quality of life. All questions were scored from 1 to 5 except for questions 16, 17, and 18 in the physical function domain, in which appropriate adjustments were made according to the scoring algorithm (questions 16 and 17 were equivalent to questions 24 and 26 in QUALEFFO-41 and were scored from 1–3, question 18 was equivalent to question 27 in QUALEFFO-41 and was scored from 1–4). For questions 16 and 17, the answers “not applicable” and “no cinema or theater within reasonable distance” were not scored. Questions 2, 3, 4, 6, 8, and 9 of the mental function domain (equivalent to questions, 33, 34, 35, 37, 39, and 40 in QUALEFFO-41) were reversed so that all scores were coded from healthiest [1] to least healthy [5]. Individual domain scores were calculated by summation of the scores of all the questions within that domain followed by linear transformation to a scale of 100, where 0 represented poor health and 100 represented good health. Floor and ceiling effects were assessed by calculating the percentage of subjects with the lowest and highest possible domain scores, respectively. This was used to confirm that the answer scale was adequate for the study population. Convergent and discriminant validity were used to assess the adequacy of score construction. Convergent validity was considered adequate if the correlation coefficient between each question with its own total domain score was  $>0.40$ . Discriminant validity was considered adequate if the correlation of the score of each question with its total domain score was higher than that of the total scores of other domains. Internal consistency was assessed using Cronbach's  $\alpha$ . The closer the Cronbach's  $\alpha$  is to 1, the higher the proportion of variance is to zero. Test–

retest repeatability was assessed by weighted kappa statistics.

Scores on SF-36 were calculated according to the standard scoring algorithm. Correlation between scores of similar domains of QUALEFFO-31C and SF-36 were analyzed. For comparison, scores were standardized such that higher scores indicated greatest impact on health (i.e., worse health). All statistical analyses were performed using SPSS for Windows version 15.0 statistical software (SPSS, Chicago, IL, USA). Receiver operating characteristics (ROC) curves were established to compare the ability of QUALEFFO and SF-36 domains as well as total and composite scores to discriminate between cases and controls over all possible cut-off values of the questionnaire. ROC curve analysis was conducted using MedCalc package version 9.3 (MedCalc, Mariakerke, Belgium).

## Results

### Subjects

For the validation study, 118 clinical fracture case–control pairs were recruited. The mean age for the clinical fracture group and control group was 63.7 (SD 6.6) and 63.4 (SD 6.8) years, respectively. The ratio of men to women was approximately 1:3. For the morphometric fracture QOL study, 68 morphometric fracture case–control quadruples (each morphometric fracture case was matched to one clinical fracture case and two controls) were included in the study. The mean age of the morphometric, clinical fracture and control subjects was 65.8 (SD 7.8), 65.4 (SD 7.4), and 65.3 (SD 7.4) years, respectively.

### Validation of the structure of QUALEFFO-31C

The results of the multi-trait analysis of the QUALEFFO-31C grouped according to three main domains—pain, physical function, and mental function—are shown in Table 1. In our study, 27 out of 31 questions demonstrated satisfactory convergent validity (i.e., item-associated domain score correlation coefficient  $>0.40$ ). The item-associated domain score correlation coefficient ranged from 0.72–0.93 for questions in the pain domain, 0.26–0.74 for the physical functioning domain, and 0.46–0.68 for the mental functioning domain. For discriminatory validity, all questions within QUALEFFO-31C demonstrated higher correlation coefficients between item-associated domain total score than item-unassociated domain total scores: pain domain, physical function domain, and mental function domain. In addition, the response rate was very high with a range of 99.2–100%. The median kappa scores for test–retest repeatability ranged from 0.65–0.85. Of the 31 questions, 23 questions (74%) had

**Table 1** Results of multi-trait analysis of QUALEFFO-31

Domain (no. of question)	Convergent validity (%)	Item convergent validity range of correlations	Discriminant validity (%)	Response rate Median (range)	Repeatability Median (range)	Internal consistency Cronbach's $\alpha$
Pain [4]	100	0.72–0.93	100	100 (99.2–100)	0.85 (0.83–0.87)	0.87
Physical function [18]	78	0.26–0.74	100	100 (99.2–100)	0.85 (0.34–0.97)	0.83
Mental function [9]	100	0.46–0.68	100	100 (99.4–100)	0.65 (0.40–0.78)	0.72

kappa values >0.7. The internal consistency ranged from 0.72–0.87 (i.e., Cronbach's  $\alpha$  >0.70).

Comparison of domain, total, and composite scores between clinical vertebral fracture cases and controls using QUALEFFO-31C and SF-36

The mean scores on QUALEFFO-31C and SF-36 for clinical vertebral fracture cases and controls are shown in Table 2. Subjects with clinical osteoporotic vertebral fractures showed significant impairment of HRQoL on the QUALEFFO-31C compared with age- and sex-matched controls. Similar results were also observed using the SF-36. All corresponding domain scores and the physical composite score showed significant impairment of QOL in clinical fracture cases relative to controls. Despite the results for the mental health domain ( $p=0.04$ ), the mental composite summary score on SF-36 did not show a significant difference between clinical fracture cases and controls ( $p=0.053$ ).

Spearman rank correlation coefficients between scores of similar domains of QUALEFFO-31C and SF-36 instruments

Spearman correlation coefficients between corresponding domains on the two scales are shown in Table 3. The correlation between corresponding domains of QUALEFFO-31C and SF-36 was good, implying that they were directly

comparable. The total QUALEFFO-31C score also correlated well with the SF-36 physical composite score ( $r^2=0.67$ ,  $p<0.001$ ).

ROC curve analysis

The discriminatory capacity of each of the domains of QUALEFFO-31C and SF-36 (as demonstrated by the area under the ROC curve) is shown in Table 4. Table 4 demonstrates that the pain domain had the highest discriminatory capacity on both instruments, followed by physical functioning, then mental functioning domains. The QUALEFFO-31C total score and SF-36 physical composite summary score had similar discriminatory capacities. The mental composite summary score of SF-36 nonetheless had poor discriminatory capacity to differentiate between the two groups of subjects. All domains, total score, and composite summary score of the QUALEFFO-31C were discriminatory between clinical fractures and controls subjects. For the SF-36, the mental composite score was not discriminatory between clinical fracture subjects and controls, but all other relevant domain and composite scores were.

Quality of life in subjects with morphometric fractures

The mean scores for various domains on QUALEFFO-31 and SF-36 for subjects with morphometric fractures and

**Table 2** Scores of QUALEFFO-31C domains in control subjects vs clinical fracture subjects

Domain	Control Mean (SD)	Clinical Mean (SD)	$p$ value
QUALEFFO			
Pain	0.9 (4.7)	33.6 (23.5)	<0.001
Physical function	6.8 (7.0)	12.2 (10.8)	<0.001
Mental function	33.3 (13.3)	37.3 (15.3)	0.035
QUALEFFO-31C total score	15.2 (7.4)	22.8 (10.3)	<0.001
QUALEFFO-31C pain + physical composite score	3.8 (4.6)	22.9 (14.6)	<0.001
SF-36			
Bodily pain	85.4 (18.7)	62.6 (22.6)	<0.001
Physical health	82.1 (17.5)	72.3 (21.4)	<0.001
Mental health	76.9 (14.3)	70.4 (18.3)	0.04
Physical composite summary score	14.2 (1.5)	13.0 (1.8)	<0.001
Mental composite summary score	27.8 (1.3)	27.4 (1.6)	0.053



**Table 3** Spearman rank correlation between scores of similar domains of QUALEFFO-31C and SF-36 instruments

QUALEFFO domain	SF-36 domain	Correlation coefficient	<i>p</i> value
Pain	Bodily pain	0.56	<0.001
Physical function	Physical functioning	0.72	<0.001
Mental function	Mental health	0.63	<0.001
QUALEFFO-31C total score	Physical composite score	0.67	<0.001
QUALEFFO-31C total score	Mental composite summary score	0.38	<0.001
QUALEFFO-31C pain + physical composite score	Physical composite summary score	0.62	<0.001

their comparison with control and clinical fracture subjects are shown in Table 5. Compared with subjects with clinical fractures, those with morphometric fractures reported a better quality of life in terms of pain and physical functioning (as seen from the respective domain scores on both instruments, the QUALEFFO-31 total score, the QUALEFFO-31 pain plus physical composite scores, and the SF-36 physical composite score). Compared with controls, subjects with morphometric fractures did not show any significant difference in HRQoL. Table 6 showed similar results as Table 5.

## Discussion

QUALEFFO is a disease-specific instrument to assess common life problems faced by patients with osteoporotic vertebral fractures. Although QUALEFFO-31C had not been previously validated in Chinese, the response rate was high in the present study (i.e., 100% for all QUALEFFO domains). QUALEFFO-31C also demonstrated good short-term reproducibility after 4 weeks and was well accepted by our population, as shown by the high kappa scores from the test–retest reproducibility analysis. The coefficient of internal consistency (Cronbach's  $\alpha$ ) of all domains was >0.70. Our results were similar to that of a previous study by Lips et al. [9] in which 26 out of 41 (i.e., 63%)

questions had kappa scores  $\geq 0.70$ . In this study, the Cronbach's  $\alpha$  for the mental functioning domain was generally lower than that of the other two domains (i.e., pain and physical function). These Cronbach's  $\alpha$  values were similar to those found in previous validation studies for both QUALEFFO-41 [9] and QUALEFFO-31 [16]. These findings suggest that QUALEFFO is easy to understand and that both controls and subjects with vertebral fractures can complete the questionnaire without major difficulties.

The original instrument (QUALEFFO-41) contained 41 questions in five domains (pain, physical functioning, mental functioning, general health perception, and social functioning). This has been reduced to 31 questions in three domains (pain, physical functioning, and mental functioning) in QUALEFFO-31. Questions from the two deleted domains were either reclassified or deleted depending on their contribution to the overall assessment. For example, questions 3 and 4 from QUALEFFO-41 were combined to make the assessment more comprehensive (i.e., producing question 3 in QUALEFFO-31) [16]. Multi-trait analysis revealed that the score construction of the new questionnaire was sufficient. This was to be expected, as the Likert scale inherited from its parent instrument was unchanged. Adequate convergent and discriminant validity was seen, indicating that questions related to the same concept had roughly equal variance, which in turn implied that weighing

**Table 4** Area under ROC in clinical fracture vs. control groups

Domain	QUALEFFO-31C (SE) <sup>a</sup>	SF-36 (SE) <sup>a</sup>	Difference in area (SE)	<i>p</i> value
Pain	0.91 (0.02)	0.79 (0.03)	0.12	<0.001
Physical function	0.69 (0.03)	0.66 (0.03)	0.03	0.249
Mental function	0.59 (0.04)	0.62 (0.04)	−0.03	0.202
QUALEFFO total score vs. SF-36 physical composite summary score	0.75 (0.03)	0.71 (0.03)	0.04	0.425
QUALEFFO total score vs. SF-36 mental composite summary score	0.75 (0.03)	0.55 (0.04)	0.20	<0.001
QUALEFFO pain and physical composite score vs. SF-36 physical composite summary score	0.91 (0.02)	0.71 (0.03)	0.20	<0.001
QUALEFFO mental score vs. SF-36 mental composite summary score	0.59 (0.03)	0.55 (0.04)	0.04	0.051

<sup>a</sup> All domain and composite score were discriminatory between clinical fracture subjects and controls except for the SF-36 mental composite summary score

**Table 5** Scores of QUALEFFO-31C domains in morphometric fracture subjects vs. clinical fracture subjects and vs. control subjects

	Morphometric ( <i>n</i> =69) Mean (SD)	Clinical ( <i>n</i> =69) Mean (SD)	Control ( <i>n</i> =133) Mean (SD)	Morphometric vs clinical fracture <i>p</i> value	Morphometric vs control <i>p</i> value
<b>QUALEFFO</b>					
Pain	0.9 (5.4)	35.2 (25.1)	0.9 (4.8)	<0.001	1.000
Physical function	8.0 (10.2)	14.0 (11.0)	6.5 (7.0)	<0.001	0.840
Mental function	34.5 (15.9)	38.6 (15.8)	33.1 (12.9)	0.289	1.000
QUALEFFO-31C total score	16.5 (9.8)	24.5 (10.6)	15.0 (7.2)	<0.001	0.777
QUALEFFO-31C pain + physical composite score	4.5 (5.9)	24.6 (15.0)	3.7 (4.6)	<0.001	1.000
<b>SF-36</b>					
Bodily pain	82.5 (22.8)	61.7 (21.2)	86.2 (18.5)	<0.001	0.721
Physical health	81.5 (17.8)	67.6 (23.0)	83.8 (16.5)	<0.001	1.000
Mental health	75.2 (15.4)	67.2 (17.9)	77.4 (14.3)	0.009	1.000
Physical composite summary score	14.1 (1.5)	12.7 (1.9)	14.3 (1.4)	<0.001	1.000
Mental composite summary score	27.7 (1.4)	27.2 (1.6)	27.8 (1.3)	0.278	1.000

Results obtained using ANOVA with Bonferroni adjustment. Subjects matched for sex, age  $\pm 3$  years

of questions was unnecessary [9]. A small number of questions in the physical function domain did not show a correlation coefficient  $>0.40$  with the physical function domain score. This was nonetheless anticipated as the study by Lips et al. [9] also noted that some of their questions within the physical function domain did not demonstrate adequate convergent validity. One of our questions in particular had a low correlation coefficient with the physical domain score (0.26). This was in accordance with the findings of a Turkish study [14] that showed similar correlation coefficients (0.27) in some of its questions in its version of QUALEFFO-41. The reason for inadequate convergent validity has not been explained in the published literature for QUALEFFO. We are uncertain about the cause of this phenomenon in our own study, but we believe that cultural and environmental factors are involved as this phenomenon tends to occur in different questions among different culture populations.

From ROC curve analyses, all domain and composite summary scores from QUALEFFO-31C were able to discriminate between clinical fracture subjects and age-matched controls. The “pain domain score” and “pain and physical composite summary score” displayed particularly strong ability to discriminate between the two study groups. When compared with the corresponding scores on the QUALEFFO-31C, the bodily pain and physical health domain scores of SF-36 showed comparable discriminatory capability. In addition, the QUALEFFO-31 total score had similar discriminatory capability to the SF-36 physical composite score. This is consistent with the results of previous studies using QUALEFFO-41 [9]. Our results showed that it had better discriminatory capability than the SF-36 mental composite score. The QUALEFFO-31 pain and physical composite score also discriminated well between clinical fracture subjects and controls, better than the SF-36 physical composite score.

**Table 6** Area under ROC in morphometric fracture vs. clinical fracture groups

Domain	QUALEFFO-31C (SE) <sup>a</sup>	SF-36 (SE) <sup>a</sup>	Difference in area (SE)	<i>p</i> value
Pain	0.90 (0.03)	0.75 (0.04)	0.15	<0.001
Physical function	0.72 (0.05)	0.72 (0.04)	0.00	0.181
Mental function	0.58 (0.05)	0.66 (0.04)	-0.08	0.148
QUALEFFO total score vs. SF-36 physical composite summary score	0.75 (0.04)	0.73 (0.04)	0.02	0.456
QUALEFFO total score vs. SF-36 mental composite summary score	0.75 (0.04)	0.57 (0.04)	0.20	<0.001
QUALEFFO pain and physical composite score vs. SF-36 physical composite summary score	0.90 (0.03)	0.73 (0.04)	0.17	<0.001
QUALEFFO mental score vs. SF-36 mental composite summary score	0.58 (0.04)	0.57 (0.04)	0.01	0.458

<sup>a</sup> All domain and composite score were discriminatory between morphometric fracture subjects and clinical fracture subjects except for the SF-36 mental composite score

Our study showed that Chinese subjects with clinical vertebral fractures suffer significant impairment in QOL compared with age and sex-matched control subjects. This was consistent with previous studies in which pain, physical, and mental functioning were all impaired. This was also apparent during our assessment using both instruments. Also consistent with previous studies was that QUALEFFO was better able to assess pain—an area known to be more seriously affected in subjects with osteoporosis. Assessment of mental functioning domains and mental composite scores had poor discriminatory power on both scales. This was again in line with the findings of the first validation study of QUALEFFO by Lips et al. [9] and subsequently confirmed by reports comparing QUALEFFO with SF-36 [19].

Further analyses demonstrated that the QUALEFFO-31 was able to discriminate well between morphometric subjects and clinical subjects in terms of pain and physical functioning. The total QUALEFFO-31 score, pain plus physical composite score were also useful discriminatory tools between these two groups. Similar results were obtained for various corresponding domains of SF-36. Like the results for clinical fractures vs. controls, the QUALEFFO-31 was superior in discriminating between clinical vs. morphometric fracture subjects in terms of pain compared with SF-36. In all other areas the two instruments had similar discriminatory ability. Results from the study also demonstrated that there was no significant difference in the quality of life between subjects with morphometric fractures and controls. Previous studies have shown that subjects with incident morphometric fractures had increased back pain and reduced physical functioning; nonetheless, those with prevalent morphometric fractures only, regardless of degree, did not show any significant results. Our results were consistent with the latter studies.

There were several limitations to the present study. First, responsiveness of QUALEFFO to incident fracture and the relationship between the number of fractures and the QUALEFFO-31 domain scores was not investigated. We had not followed the patients over a period of time for fracture outcome to evaluate whether there is any change in the scores. Previous studies reveal that QUALEFFO-41 did not have a significant relationship with the number of fractures; nonetheless, the preliminary study of QUALEFFO-31 showed ability to differentiate between those with only one fracture versus those with more than one vertebral fracture. As this study lacks long-term follow-up, and to avoid erroneous conclusions from subject's recall of vertebral fracture history we have not attempted to determine whether a similar relationship exists for QUALEFFO-31C compared with the English version. A longitudinal study would be helpful to validate this instrument in our population. Second, the main aim of our study was to validate the Chinese version of

QUALEFFO-31. Cognitive debriefing had not been carried out to ascertain the understanding, interpretation, and relevance of the content of the questionnaire in our culture. It is likely that some items are not fully applicable to our population, which might explain the low Cronbach's alpha in some domains. Besides, the QOL of incident fractures (both morphometric and clinical) and the QOL of subjects after medical/surgical treatment were not investigated in this study. These are interesting topics that may prove helpful if investigated.

## Conclusions

In conclusion, quality of life is impaired in subjects with vertebral fractures due to osteoporosis when compared with controls. This study also demonstrated that the Chinese version of QUALEFFO-31 is a reliable, repeatable, consistent, coherent tool that discriminates well between clinical fracture subjects and controls, particularly in terms of pain and physical function. However, the Cronbach's  $\alpha$  results suggest that further refinement in specific areas of the Chinese version of QUALEFFO-31 is required, especially in the mental functioning domain. Taken together, this new shortened form of QUALEFFO reduces redundancy of questions and effectively assesses health-related quality of life for subjects with clinical vertebral fractures. The findings will be useful for the determination of cost-effective threshold cut-off for Chinese in the future.

**Acknowledgments** The authors wish to thank the nursing and technical staff of the Osteoporosis Centre, Department of Medicine, Queen Mary Hospital for their help in carrying out this project. This study was supported by the Bone Health Fund of the Hong Kong University Foundation and Osteoporosis Research Fund of the University of Hong Kong.

**Disclosures** None.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Kung AW, Yates S, Wong V (2007) Changing epidemiology of osteoporotic hip fracture rates in Hong Kong. *Arch Osteoporos* 2:53–58
2. Randell A, Sambrook PN, Nguyen TV, Lapsley H, Jones G, Kelly PJ, Eisman JA (1995) Direct clinical and welfare costs of osteoporotic fractures in elderly men and women. *Osteoporos Int* 5:427–432
3. Lau EM, Chan HH, Woo J, Lin F, Black D, Nevitt M, Leung PC (1996) Normal ranges for vertebral height ratios and prevalence of



- vertebral fracture in Hong Kong Chinese: a comparison with American Caucasians. *J Bone Miner Res* 11:1364–1368
4. Ross PD, Fujiwara S, Huang C, Davis JW, Epstein RS, Wasnich RD, Kodama K, Melton LJ 3rd (1995) Vertebral fracture prevalence in women in Hiroshima compared to Caucasians or Japanese in the US. *Int J Epidemiol* 24:1171–1177
  5. Iki M, Kagamimori S, Kagawa Y, Matsuzaki T, Yoneshima H, Marumo F (2001) Bone mineral density of the spine, hip and distal forearm in representative samples of the Japanese female population: Japanese Population-Based Osteoporosis (JPOS) Study. *Osteoporos Int* 12:529–537
  6. Lau EM, Chan YH, Chan M, Woo J, Griffith J, Chan HH, Leung PC (2000) Vertebral deformity in Chinese men: prevalence, risk factors, bone mineral density, and body composition measurements. *Calcif Tissue Int* 66:47–52
  7. Cook DJ, Guyatt GH, Adachi JD, Epstein RS, Juniper EF, Austin PA, Clifton J, Rosen CJ, Kessenich CR, Stock JL, Overdorf J, Miller PD, Erickson AL, Mc CMR, McClung BL, Griffith LE, Ioannidis G (1999) Development and validation of the mini-osteoporosis quality of life questionnaire (OQLQ) in osteoporotic women with back pain due to vertebral fractures. Osteoporosis Quality of Life Study Group. *Osteoporos Int* 10:207–213
  8. Lau EM, Woo J, Chan H, Chan MK, Griffith J, Chan YH, Leung PC (1998) The health consequences of vertebral deformity in elderly Chinese men and women. *Calcif Tissue Int* 63:1–4
  9. Lips P, Cooper C, Agnusdei D, Caulin F, Egger P, Johnell O, Kanis JA, Kellingray S, Leplege A, Liberman UA, McCloskey E, Minne H, Reeve J, Reginster JY, Scholz M, Todd C, de Vernejoul MC, Wiklund I (1999) Quality of life in patients with vertebral fractures: validation of the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO). Working Party for Quality of Life of the European Foundation for Osteoporosis. *Osteoporos Int* 10:150–160
  10. Pluijm SM, Tromp AM, Smit JH, Deeg DJ, Lips P (2000) Consequences of vertebral deformities in older men and women. *J Bone Miner Res* 15:1564–1572
  11. Ross PD (1997) Clinical consequences of vertebral fractures. *Am J Med* 103:30S–42S, discussion 42S–43S
  12. Silverman SL (1992) The clinical consequences of vertebral compression fracture. *Bone* 13(Suppl 2):S27–S31
  13. Cockerill W, Lunt M, Silman AJ, Cooper C, Lips P, Bhalla AK, Cannata JB, Eastell R, Felsenberg D, Gennari C, Johnell O, Kanis JA, Kiss C, Masaryk P, Naves M, Poor G, Raspe H, Reid DM, Reeve J, Stepan J, Todd C, Woolf AD, O'Neill TW (2004) Health-related quality of life and radiographic vertebral fracture. *Osteoporos Int* 15:113–119
  14. Kocyigit H, Gulseren S, Erol A, Hizli N, Memis A (2003) The reliability and validity of the Turkish version of Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO). *Clin Rheumatol* 22:18–23
  15. Ramirez Perez E, Clark P, Wachter NH, Cardiel MH, Pilar Diez Garcia M del (2008) Cultural adaptation and validation of the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO) in a Mexican population. *Clin Rheumatol* 27:151–161
  16. van Schoor NM, Knol DL, Glas CA, Ostelo RW, Leplege A, Cooper C, Johnell O, Lips P (2006) Development of the QUALEFFO-31, an osteoporosis-specific quality-of-life questionnaire. *Osteoporos Int* 17:543–551
  17. Genant HK, Wu CY, van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res* 8:1137–1148
  18. Lam CL, Gandek B, Ren XS, Chan MS (1998) Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. *J Clin Epidemiol* 51:1139–1147
  19. Lips P, van Schoor NM (2005) Quality of life in patients with osteoporosis. *Osteoporos Int* 16:447–455