RESEARCH ARTICLE

Reliability and Validity of an Iranian Version of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for Patients with Multiple Myeloma: the EORTC QLQ-MY20

Ahmad Ahmadzadeh[1], Mir Saeed Yekaninejad[2], Mohsen Saffari[3], Amir H Pakpour[4]*, Neil K Aaronson[5]

Abstract

Background: Reliable and validated instruments are needed in order to study the quality of life in myeloma patients. This study aimed to translate and explore the psychometric properties of the European Organisation for Research and Treatment of Cancer (EORTC) myeloma module (QLQ-MY20) in Iranian patients. Materials and Methods: Two hundred and fifteen patients with multiple myeloma (MM) were recruited from Imam Khomeini Hospital, Tehran. A standard forward-backward translation procedure was implemented. Participating patients were asked to complete the EORTC QLQ-C30 and the QLQ-MY20 three times, at study entry, after two weeks, and again after three months. Data were tested for the range of measurement, internal consistency, test-retest reliability, known group comparison, responsiveness and factor structure. Results: Mean age of the patients was 60.7 years. No floor and ceiling effects were seen for the QLQ-MY20. Cronbach's α was greater than 0.80 for all three multi-item scales (ranging from 0.82 to 0.93). All four scales had test-retest reliability of 0.85 or greater. Results of the confirmatory factor analysis that the hypothesized 3-scale measurement model of the QLQ-MY20. Moreover, the Persian version for the QLQ-MY20 differentiated between subgroups of the patients in terms of beta-2 microglobulin, fracture and performance status. The responsiveness of the QLQ-MY20 to change over time was confirmed within 3 months. Conclusions: the results of our study indicate that our Iranian version of the QLQ-MY20 is a feasible, reliable and valid questionnaire for assessing the condition-specific quality of life of patients with MM.

Keywords: Quality of life - validation - EORTC QLQ-MY20 - cancer - outcome assessment - psychometrics - Iran

Introduction

Multiple myeloma (MM) is a malignant plasma-cell proliferative disorder. It is the second most common hematologic malignancy in the United States (Jemal et al., 2004). The overall estimated annual incidence of MM is 86,000 cases (Alexanian and Dimopoulos, 1994). In the United States, MM is estimated to account for about 11,000 deaths each year (Kyle and Rajkumar, 2004). The etiology of MM is not well understood, but lifestyle, occupational and environmental risk factors have been linked to the disease, as have age and male gender (Alexanian and Dimopoulos, 1994). The average 5-year survival of MM is estimated to be 33% (Alexanian and Dimopoulos, 1994).

Despite significant improvement in treatment and management of MM, it still remains incurable (Kyle and Rajkumar, 2004). Therefore, palliating symptoms and prolonging patient's life are considered as the major aims of MM treatment (Kyle and Rajkumar, 2004). However, MM treatment is associated with a number of side effect and toxicities that can increase the symptom burden associated with MM (e.g., bone pain, nausea, vomiting, fractures, hypercalcaemia, and fatigue) (Johnsen et al., 2009). These symptoms can, in turn, have a negative impact on the patient's daily functioning and quality of life.

Health-related quality of life (HRQOL) is a potentially important endpoint in studies in MM, providing relevant data that can help physicians and patients choose the

[1]Research Centre of Cancer, Shafa Hospital, Ahvaz Jundishapur University of Medical Science, Ahvaz, [2]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, [3]Health Research Center, Baqiyatallah University of Medical Sciences, Tehran, [4]Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Qazvin, Iran, [5]Division of Psychosocial Research & Epidemiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands *For correspondence: Pakpour_Amir@yahoo.com

most appropriate treatment. Although HRQOL has been assessed in a number of clinical studies in MM, a recent review concluded that there is as yet no consensus on how it should be assessed (Kvam et al., 2009). The most commonly used HRQOL questionnaire in MM research is the EORTC QLQ-C30 (Aaronson et al., 1993). Although the QLQ-C30 provides valuable information on generic aspects of HRQOL, it does not cover a range of MM-specific issues (Osborne et al., 2012). For this reason, the EORTC developed a supplemental, myeloma-specific module (the QLQ-MY24), which has validated in studies including patients from Europe and the United States (Kontodimopoulos et al., 2012). However, the QLQ-MY24 has not been used or evaluated previously in the Iranian setting. The aim of the study was therefore to translate and evaluate psychometrically the Iranian version of the QLQ-MY24.

Materials and Methods

Study sample

The study was conducted from July, 2013 to March, 2014. The study sample consisted of 215 patients with multiple myeloma referred for treatment to the Imam Khomeini Hospital, a large university hospital in Tehran. Patients were eligible for the study if they were 18 years or older, had been diagnosed with multiple myeloma at least one month prior to study entry, were able to read and speak Persian/Farsi. Patients who had a life expectancy of less than 3 months or had serious cognitive problems (as assessed by the Mini-Mental State Examination) were excluded from the study. The study protocol was approved by the ethics committee of the Qazvin University of Medical Sciences. All patients provided written informed consent to participate in the study.

Data collection and measures

Participating patients were asked to complete the EORTC QLQ-C30 and the QLQ-MY20 three times, at study entry, at two weeks, and again at three months.

The EORTC-C30

The EORTC-C30 (version 3.0) is a cancer-specific measure of quality of life in cancer patients. It consists of 30 items organized into nine multi-item scales and six single items. The EORTC-C30 has been translated into a large number of languages, including Persian. The Persian version of the EORTC-C30 has been found to be valid and reliable when used with Iranian patients with cancer (Montazeri et al., 1999).

The EORTC QLQ-MY20

The EORTC QLQ-MY20 has been developed as a supplement to the QLQ-C30 specifically for use among patients with multiple myeloma. It consists of 20 questions grouped into three scales assessing future perspective (3 items), disease symptoms (6 items), side effects of treatment (10 items), and a single item on body image. Responses are scored on a four-point scale, ranging from "not at all" (1) to "very much" (4). Scores are linearly transformed to a 0-100 scale. (Cocks et al., 2007). Higher

scores on the body image and future perspective scales represent better outcomes, and higher scores on the symptoms and side effect scales poorer outcomes.

Translation procedure

Two translators bilingual in Persian and English and with a medical background translated independently the original English version of the QLQ-MY20 into Persian. Subsequently, the coordinator of the translation process compared these translated versions and reconciled disparities to create a single, merged forward translation. This forward Persian translation of the QLQ-MY20 was translated back into English by two native English speakers with a high level of fluency in Persian. The coordinator compared the English back translations with the original questionnaire and resolved differences in a session with the translators. The provisional translation was then piloted on 14 multiple myeloma patients (average age 58.12 [SD=9.14], 8 female and 6 male). The average time to complete the QLQ-MY20 was less than 6 minutes and the questionnaire was well understood and acceptable to patients.

Statistical analysis

We examined the QLQ-MY20 score distributions, including the percentage of scores at the extremes of the scale. Floor or ceiling effect were considered present if more than 15% of respondents reported the highest or lowest possible score (Terwee et al., 2007).

We used Cronbach's coefficient alpha (α) to assess internal consistency reliability, and considered α's equal to or greater than 0.7 to be acceptable. We evaluated test–retest reliability by calculating the intraclass Correlation Coefficient (ICC) (Rosner, 1995).

We performed a confirmatory factor analysis (CFA) to examine the hypothesized 3-scale measurement model of the QLQ-MY20; future perspective, disease symptoms and side effects of treatment. The model fit was evaluated using Chi-square (χ^2), the adjusted goodness of fit index (AGFI), χ^2/df ratio, the standardized root mean square residual (SRMR) , the Root Mean Square Error of Approximation (RMSEA), the Non-Normed Fit Index (NNFI), the Comparative Fit Index (CFI), and the parsimonious normed fit index (PNFI). The thresholds of goodness-of-fit indices were as follow; a non-significant chi square (p>0.05), χ^2/df ratio of 3:1 or smaller, a RMSEA less than 0.08, a CFI 0.90 or greater, a NNFI 0.90 or greater, an AGFI 0.90 or greater and a SRMR of 0.08 or less generally were considered a good fit [17].

We performed known-group validity testing, using the Wilcoxon Mann-Whitney statistic, to determine how well the QLQ-MY20 is able to discriminate between subgroups of patients formed on the basis of performance status, fractures and b-2 microglobulin level. It was hypothesized that patients with lower performance status, any fracture and high b-2 microglobulin level ($\geq$4 mg/l) would report poorer scores on the questionnaire (Wisloff and Hjorth, 1997; Cocks et al., 2007). All p-values were adjusted for multiple comparisons according to the Benjamini–Hochberg procedure (Benjamini et al., 2006).

Finally, we assessed the responsiveness of the QLQ-

MY20 to change over time in patients who responded to the treatment (i.e. complete response [CR], partial response [PR] or near complete response [NCR]). Change in QLQ-MY20 scores at baseline and three months after treatment in patients who responded to the treatment was computed using Wilcoxon rank sum test. Standardized response mean (SRM), an effect size index, was calculated to estimate the magnitude of the change over time (Liang et al., 1990).

Data analyses were conducted with the SPSS 18.0 (SPSS, Chicago, IL, USA) and LISREL 8.80.

Results

Demographic and clinical characteristics of the samples are summarized in Table 1. Mean age of the patients was 60.7 years (SD=14.1). Almost all patients (201, 93.5%) completed the questionnaire in less than 6

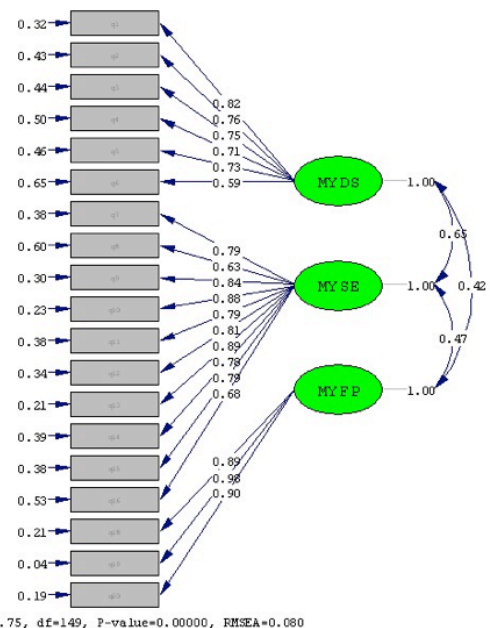**Table 1. Characteristics of the Subjects**

Age (years) Mean (SD)	60.71 (14.12)
Sex	
Male	120 (55.8%)
Female	95 (44.2%)
Educational status	
Illiterate	60 (27.9%)
Primary school	46 (21.4%)
Middle school	43 (20.0%)
Secondary school	45 (21.0%)
College	21 (9.8%)
Marital status	
Single	28 (13.0%)
Married	160 (74.5%)
Widowed/divorced	27 (12.5%)
Occupational status	
Employee	76 (35.4%)
Unemployed	139 (64.7%)
World Health Organisation performance status (PS)	
0	22 (10.2%)
1	102 (47.4%)
2	65 (30.2%)
3	17 (7.9%)
4	9 (4.2%)
Disease status	
Newly diagnosed	179 (83.3%)
Relapsed/refractory	36 (16.7%)
Fractures present	
Yes	73 (34.0%)
No	142 (66.0%)
β-2 Microglobulin (mg/L)	
<4	81 (37.7%)
7-Apr	47 (21.9%)
≥8	34 (15.8%)
Missing	53 (24.6%)
Transplant received	
Yes	52 (24.2%)
No	163 (75.8%)
Receiving bisphosphonate	
Yes	174 (80.9%)
No	41 (19.1%)
Chemotherapy	
Conventional	147 (68.4%)
High dose	68 (31.6%)

minutes, with a mean of 5.12 (SD=2.24). Floor (<12%) and ceiling effects (<14%) were low for all scales, varying between 4% and 13% (Table 2). Cronbach's α was greater than 0.80 for all three multi-item scales (ranging from 0.82 to 0.93) (Table 2). All four scales (including the single item body image measure) had test-retest reliability of 0.85 or greater. The disease symptoms scale had the highest test-retest reliability (ICC=0.95).

Results of the CFA also are presented in Figure 1. They indicate that the three-factor, 19-item model converged and had an acceptable but marginal fit (χ^2=259.75, degree of freedom=149, p<0.0001, RMSEA=0.080, SRMR=0.074, AGFI=0.96, CFI=0.97, NNFI=0.97, and PNFI=0.82). All scale items loaded significantly on related constructs and the standardized factor loadings were all above the minimum recommended value of 0.4. The correlations between the latent factors were significant and ranged from 0.42 to 0.65. Factor loadings ranged from 0.59 to 0.98.

As hypothesized, patients with equal or higher levels of beta 2-microglobulin reported lower levels of functioning and higher levels of symptoms (Table 3). The effect sizes for the QLQ-MY20 subscales were all in the medium to large range (range 0.32 to 1.73). Low performance status was related significantly to lower levels of functioning and higher levels of symptoms, with effect sizes ranging from 0.38 to 0.46. More than 65% of the patients reported having a fracture (n=142). Scores on both the disease symptom and treatment side effects scales of the QLQ-MY20 were significantly higher in patients with a fracture than those without any fracture at baseline (effect size ranged from 0.0.32 to 0.53). Conversely, patients without any fracture at baseline reported higher levels of functioning (Table 3).

The responsiveness of the QLQ-MY20 to change over time (i.e., pre- versus post-treatment) was assessed for those patients who achieved at least a PR (Table 4). There

Chi-Square=259.75, df=149, P-value=0.00000, RMSEA=0.080

Figure 1. Standardized Estimated Factor-item Loadings, Error Variances, and Covariance for the Final Model of the QLQ-MY20

Table 2. Summary statistics for the QLQ-MY20 at baseline (N=215)

	Number of forms	Mean	SD	Floor N (%)	Ceiling (%)	Cronbach's alpha	Normality
Future perspective	209	40.68	29.06	24(11.2%)	16 (7.4%)	0.816	0.02
Body image	215	77.57	37.55	17 (7.9%)	29 (13.5%)	-	0.004
Disease symptoms	214	36.46	20.64	11 (5.1%)	22 (10.2%)	0.853	0.01
Side effects of treatment	215	29.59	16.54	9 (4.2%)	18 (8.4%)	0.938	0.008

Table 3. Known-groups Validity the QLQ-MY20 Questionnaire

	beta-2 microglobulin			performance status			fractures		
	<4 mg/l (N=81)	≥4 mg/l (N=81)	Effect size	Low (N=124)	High (N=91)	Effect size	Present (N=73)	Absent (N=142)	Effect size
Future perspective	62.22 (24.69)	22.11 (12.45)	2.05	34.89 (20.83)	46.18 (32.37)	0.43	31.41 (21.52)	77.77 (38.49)	1.37
Body image	88.37 (31.84)	68.14 (22.80)	0.73	62.62 (33.40)	78.84 (37.32)	0.46	59.91 (24.72)	79.44 (37.46)	0.58
Disease symptoms	36.78 (21.48)	68.22 (26.58)	1.3	44.80 (27.86)	34.31 (22.53)	0.41	43.66 (26.74)	33.83 (32.67)	0.32
Side effects of treatment	19.33 (9.81)	31.87 (13.78)	1.05	33.02 (13.31)	28.41 (10.64)	0.38	32.51 (14.12)	26.06 (11.02)	0.53

*WHO performance status: High: scores 0/1. Low: score 2, 3, 4 (baseline), [†] Statistically significant for WHOPS, [‡] Statistically significant for response to treatment

Table 4. Responsiveness of the MY20 for Patients who Responded to the Treatment (n=105)

	Baseline Mean (SD)	Follow-up	SRM* 0-3 months	P value[a]
Future perspective	39.64 (29.15)	44.84 (20.51)	0.3	0.002
Body image	71.84 (38.23)	63.78 (31.39)	0.39	0.0003
Disease symptoms	38.23 (20.28)	30.22 (19.41)	0.23	0.0001
Side effects of treatment	29.21 (13.86)	31.97 (18.45)	0.19	0.012

* SRM standardized response mean, [a]Wilcoxon signed-rank test

were a total of 105 patients (49%) at follow-up who had either a CR, PR or NCR to their treatment. The patients reported significantly more side effects of treatment (p = 0.01), and poorer body image (p < 0.000) over time. Conversely, disease symptoms (p <0.000) and future perspective (p = 0.002) improved significantly over time. The effect sizes were relatively modest (ranging from , 0.19 for the treatment side effects scale to 0.39 for body image.

Discussion

The objective of this study was to examine the psychometric properties of the Iranian version of the QLQ-MY20 among patients with MM. The results show that the Iranian version of the QLQ-MY20 is acceptable to patients (with a response rate of more than 95% and only very minor amounts of missing data (<2%) for individual items). No significant floor or ceiling effects were observed.

All of the QLQ-MY20 scales exceeded the recommended minimum alpha coefficient standard of 0.70 for group assessment. In fact, the alpha coefficients observed in the current study were somewhat higher than

those reported in the original English and Greek versions (Cocks et al., 2007; Kontodimopoulos et al., 2012).

The test-retest reliability for the Iranian version of the QLQ-MY20 was investigated in patients who were clinically stable over a two-week interval. The ICCs were high for all scales. To our knowledge, no previous studies have reported on the test-retest reliability of this questionnaire. Future studies should examine reproducibility of the QLQ-MY20 across English-speaking and multi-national samples.

The results of CFA in this study yielded a three structure of the Iranian version of the QLQ-MY20. The CFA had a good internal consistency and the empirically derived factor structure is supported by Cocks et al [10]. Most of the items loaded as expected and the factor loadings met the minimum recommended standard of 0.4 value recommended by Hair et al. (2010). To the best of our knowledge, this the first study to explore factor structure of the QLQ-MY20. We would recommend that future studies assess the factor structure of the QLQ-MY20 across different cultures and languages.

Known group validity was used to assess the ability of the QLQ-MY20 to discriminate between subgroups of the patients in terms of beta-2 microglobulin, fracture

and performance status. The results indicated that patients with lower levels of -2 microglobulin (>4 mg/l) reported better quality of life than those with higher level (≥4mg/l). In accordance with the present results, previous studies have demonstrated that low level of beta 2-microglobulin in serum was associated with response to treatment and survival in patients with MM (Greipp et al., 1993; Attal et al., 1996). Bone health is significantly affected by cancer therapy. Over 80% patients with MM experience skeletal complications such as bone loss and pathologic fracture (Sonmez et al., 2008). Bone loss places cancer patients at risk of fracture and thus compromised quality of life (Guise, 2006). The present findings are consistent with other studies that have found that patients without fractures and with high performance status report better quality of life (Cocks et al., 2007).

In conclusion, the results of our study indicate that the Iranian version of the QLQ-MY20 is a feasible, reliable and valid questionnaire for assessing the condition-specific quality of life of patients with MM. The QLQ-MY20, in combination with the QLQ-C30, provides a comprehensive picture of the impact of MM and its treatment on the HRQOL of patients.

References

Aaronson NK, Ahmedzai S, Bergman B, et al (1993). The european organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*, **85**, 365-76.

Alexanian R, Dimopoulos M (1994). The treatment of multiple myeloma. *N Engl J Med*, **330**, 484-9.

Attal M, Harousseau JL, Stoppa AM, et al (1996). A prospective, randomized trial of autologous bone marrow transplantation and chemotherapy in multiple myeloma. intergroupe francais du myelome. *N Engl J Med*, **335**, 91-7.

Benjamini Y, Krieger AM, Yekutieli D (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, **93**, 491-507.

Cocks K, Cohen D, Wisloff F, et al (2007). An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-MY20) in assessing the quality of life of patients with multiple myeloma. *Eur J Cancer*, **43**, 1670-8.

Greipp PR, Lust JA, O'Fallon WM, et al (1993). Plasma cell labeling index and beta 2-microglobulin predict survival independent of thymidine kinase and C-reactive protein in multiple myeloma. *Blood*, 81, 3382-7.

Guise TA (2006). Bone loss and fracture risk associated with cancer therapy. *Oncologist*, **11**, 1121-31.

Hair JF, Black WC, Babin BJ, et al (2010). Multivariate data analysis: a global prospective Boston, Pearson Education.

Jemal A, Tiwari RC, Murray T, et al (2004). Cancer statistics, 2004. *CA Cancer J Clin*, **54**, 8-29.

Johnsen AT, Tholstrup D, Petersen MA, et al (2009). Health related quality of life in a nationally representative sample of haematological patients. *Eur J Haematol*, **83**, 139-48.

Kontodimopoulos N, Samartzis A, Papadopoulos AA, et al (2012). Reliability and validity of the Greek QLQ-C30 and QLQ-MY20 for measuring quality of life in patients with multiple myeloma. *Scientific World J*, **2012**, 842867.

Kvam AK, Fayers P, Hjermstad M, et al (2009). Health-related quality of life assessment in randomised controlled trials in multiple myeloma: a critical review of methodology and impact on treatment recommendations. *Eur J Haematol*, **83**, 279-89.

Kyle RA, Rajkumar SV (2004). Multiple myeloma. *N Engl J Med*, **351**, 1860-73.

Liang MH, Fossel AH, Larson MG (1990). Comparisons of five health status instruments for orthopedic evaluation. *Med Care*, **28**, 632-42.

Montazeri A, Harirchi I, Vahdani M, et al (1999). The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30): translation and validation study of the Iranian version. *Support Care Cancer*, **7**, 400-6.

Osborne TR, Ramsenthaler C, Siegert RJ, et al (2012). What issues matter most to people with multiple myeloma and how well are we measuring them? A systematic review of quality of life tools. *Eur J Haematol*, **89**, 437-57.

Rosner B (1995). Fundamentals of Biostatitistcs, Belmont, CA, Duxbury Press.

Sonmez M, Akagun T, Topbas M, et al (2008). Effect of pathologic fractures on survival in multiple myeloma patients: a case control study. *J Exp Clin Cancer Res*, **27**, 11.

Terwee CB, Bot SD, de Boer MR, et al (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*, **60**, 34-42.

Wisloff F, Hjorth M (1997). Health-related quality of life assessed before and during chemotherapy predicts for survival in multiple myeloma. nordic myeloma study group. *Br J Haematol*, **97**, 29-37.