available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

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

K. Cocks[a,], D. Cohen[a], F. Wisløff[b], O. Sezer[c], S. Lee[d], E. Hippe[e], P. Gimsing[f], I. Turesson[g], R. Hajek[h], A. Smith[i], L. Graham[a], A. Phillips[a], M. Stead[j], G. Velikova[k], J. Brown[a], on behalf of EORTC Quality of Life Group*

[a]*Clinical Trials Research Unit, University of Leeds, 17 Springfield Mount, Leeds LS2 9NG, UK*
[b]*Department of Haematology, Ulleval University Hospital and Faculty of Medicine, University of Oslo, Norway*
[c]*Department of Hematology and Oncology, University Hospital Charité Berlin, Berlin, Germany*
[d]*Dana-Farber Cancer Institute, Boston, USA*
[e]*Department of Haematology, Herlev Hospital, Herlev, Denmark*
[f]*Department of Haematology, Rigshospitalet, Copenhagen, Denmark*
[g]*Malmö University Hospital, Malmö Sweden*
[h]*Department of Internal Medicine and Hemato-oncology, Faculty Hospital Brno and Faculty of Medicine MU in Brno, Czech Republic*
[i]*Southampton General Hospital, Southampton, UK*
[j]*UK Clinical Research Network Co-ordinating Centre, UK*
[k]*Cancer Research UK Clinical Centre, St. James University Hospital, Leeds, UK*

ARTICLE INFO

Article history:
Received 15 February 2007
Received in revised form 20 April 2007
Accepted 23 April 2007
Available online 15 June 2007

Keywords:
Quality of life
Multiple myeloma
Validation
Questionnaires
QLQ-C30
QLQ-MY24
QLQ-MY20

ABSTRACT

Aim: To test the reliability, validity and sensitivity of the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-MY24 questionnaire, designed to assess the quality of life of myeloma patients with the QLQ-C30.

Methods: The study was carried out through the EORTC Quality of Life Group using clinical trials in seven countries. All trials used the QLQ-C30 and QLQ-MY24 at baseline and a follow-up timepoint.

Results: Two hundred and forty patients participated. The questionnaires were acceptable to patients. The hypothesised scale structure (disease symptoms, side-effects, body image and future perspective) was confirmed by multi-trait scaling, internal consistency and correlation analysis. Most scales demonstrated sensitivity to change and discriminated between clinically different patients. The social support scale (4 items) was removed due to observed ceiling effects.

Conclusion: The final questionnaire contains 20 items, QLQ-MY20, and is a reliable and valid instrument recommended for use with the QLQ-C30 in myeloma patients.

* *Corresponding author:* Tel.: +44 113 3431475; fax: +44 113 3431471.
E-mail address: k.cocks@leeds.ac.uk (K. Cocks).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.04.022

1. Introduction

Multiple myeloma is the second most common haematological malignancy.[1] Myeloma is more common in elderly patients and is rare in patients under the age of 40. Approximately 40% of patients survive 5 years with a median survival of 3–5 years.[2,3] The full clinical picture of myeloma is a complex of bone destruction, leading to bone pain, fractures and hypercalcaemia, recurrent bacterial infections due to immune deficiency, marrow failure leading to anaemia and less commonly thrombocytopenia, renal failure and amyloidosis. These symptoms may cause pronounced distress and influence different aspects of every day life of the patients. Patients also suffer side-effects of the palliative life-prolonging systemic treatment based on more or less intensive chemotherapy and high dose dexamethasone. Measuring quality of life in these patients is therefore of great importance.

The EORTC QLQ-C30 is a psychometrically robust, cross-culturally accepted questionnaire designed to be used to assess quality of life in a variety of cancer patients. It has been validated for use in myeloma patients.[4,5] The EORTC myeloma module (QLQ-MY24) has been designed to use alongside this questionnaire to address issues of more relevance to myeloma patients. The module has been developed according to EORTC Quality of Life (QOL) Group Development Guidelines[6] and translated into 11 languages following EORTC QOL Group Translation Guidelines.[7] The early phases of development of the module have been described elsewhere.[8]

This international field study on behalf of the EORTC QOL Group aimed to determine the reliability, validity and sensitivity of the QLQ-MY24 and confirms the hypothesised scaling structure of the module.

2. Patients and methods

2.1. Patients

Patients were recruited into the field study from October 2002 to July 2004. The study was coordinated by the Clinical Trials Research Unit, University of Leeds, UK. Patients were recruited prospectively from ongoing or new clinical trials at the time of the study rather than setting up a separate questionnaire validation study. This approach has also been utilised in validation of some of the Functional Assessment of Chronic Illness Therapy (FACIT) questionnaires. Trials worldwide were considered for participation. Potential protocols were reviewed by the project team and chosen to provide a sample with a wide range of treatments and disease characteristics for the purposes of the study. Eligible studies were those assessing quality of life using the EORTC QLQ-C30 and QLQ-MY24 at baseline and at a second time point (during or on completion of trial treatment). Trials had to be for newly diagnosed or relapsed/refractory patients with a minimum prognosis of 3 months. Maintenance therapy trials were not eligible. Patients had to be mentally fit and understand the language of the questionnaire. Appropriate ethics approval was sought for participants from each trial, ensuring permission was given to transfer data for field study purposes.

2.2. Questionnaires

The EORTC QLQ-C30 and the proposed module QLQ-MY24 were issued to patients at the baseline for each trial (randomisation or registration) according to the procedures for that trial. A panel reviewed each protocol and identified an appropriate follow-up timepoint for the field study QOL assessment, this had to be during treatment or within 4 weeks of completion of treatment and was 3 or 6 months from randomisation for all protocols.

The EORTC QLQ-MY24 contained 24 items assessing aspects of QOL specific to multiple myeloma. A 'not at all', 'a little', 'quite a bit' or 'very much' response scale is used to rate each item. The hypothesised scale structure consists of four scales (two symptom scales: disease symptoms, side-effects of treatment, one function scale: future perspective and a social support scale) and one single item: body image. All scores are linearly transformed to a 0–100 scale. A higher score on the symptom scales reflects a higher level of symptoms, whereas a higher score on future perspective, social support and body image indicates good functioning or support.

A standard EORTC debriefing questionnaire[6] was used to record the time required to complete the questionnaires, whether there are items that are not covered and whether questionnaire items were considered confusing/difficult to answer or upsetting.

2.3. Statistical methods

A sample size of 240 was calculated using previous field-testing modules, where 5–10 patients per questionnaire item were required. Due to potential patient attrition before the follow-up timepoint, it was estimated that 343 patients would be required in total to achieve 240 patients with both QOL assessments completed in the correct timeframes. The aim was to recruit a minimum of 30 patients from each participating country to achieve adequate representation from each.

Analyses were carried out using SAS.[9] Bonferroni's adjustment was used to account for multiple testing. Patient characteristics, compliance and missing data were summarised by trial and country. Bar charts (not shown) were produced to show the percentages of each response for individual items, and the distribution of the responses were summarised.

2.3.1. Reliability

Internal consistency was assessed using the Cronbach's alpha coefficient.[10] As recommended by Nunnally,[11] estimates of 0.7 or greater were considered acceptable for the psychometric scales.

2.3.2. Construct validity

Multi-trait scaling analysis[12] was used to confirm that items are included in the scale they most strongly correlate with and that the proposed scale structure is consistent with the data patterns. Evidence of convergent validity was defined as a correlation of greater than 0.4 between an item and its hypothesised scale (corrected for overlap[13]). Discriminant validity compares the magnitude of the correlation of an item

with its hypothesised scale to that with other scales. A 'scaling success' was where the correlation between an item and its hypothesised scale was significantly higher, at least two times the standard error, than with another scale.

Correlations between all scales from both questionnaires were examined. Scales which are conceptually related were expected to correlate with each other with a coefficient of greater than 0.4.

2.3.3. *Clinical validity*

Known-group comparisons were made to assess the extent to which questionnaire scores were able to discriminate between subgroups of patients known to differ in terms of clinical status. Clinical variables used for this comparison were performance status (0/1/2 versus 3/4), fractures (present versus absent at baseline), β-2 microglobulin level (<4 versus ⩾4 mg/l), disease (newly diagnosed vs relapsed) and chemotherapy intensity (conventional-dose versus high-dose). Group differences were assessed for significance using the Wilcoxon rank sum test.

2.3.4. *Responsiveness to change*

The availability of questionnaires prior to the start of treatment and during or on completion of treatment allows a preliminary evaluation of the responsiveness of the questionnaires to changes in health status over time. A patient's health status is thought to have changed if they respond to treatment (complete response (CR), partial response (PR) or near complete response (NCR)). Differences between baseline and follow-up assessments for the patients who responded to treatment were assessed for statistical significance using the Wilcoxon rank sum test.

3. Results

3.1. *Patient characteristics*

A total of 477 patients from 10 trials in seven countries were registered. A high proportion of patients (54%) were from the UK due to a large myeloma trial recruiting for the duration of this field study. Sociodemographic data, clinical characteristics and treatment details of patients completing questionnaires within the required timeframes are summarised in Table 1. The final sample consisted of 225 patients with newly diagnosed disease and only 15 with relapsed or refractory disease.

3.2. *Patient compliance and debriefing questions*

To be included in the analysis, a patient must have returned both a baseline questionnaire completed before study treatment, and a follow-up questionnaire completed within the specified time frame.

Of the 477 patients registered, 343 (72%) returned a baseline questionnaire within the required timeframe. Twenty-eight of these died before follow-up, and 240 (76% of those available) returned a follow-up questionnaire within an acceptable timeframe. Patients without both questionnaires did not differ significantly from those who complied in terms

Table 1 – Sociodemographic data, clinical characteristics and treatment details

	Patients with baseline and follow-up questionnaires (number of patients = 240)
Sociodemographic data	
Country	
Czech Republic	16 (6.6%)
Denmark	19 (7.9%)
Germany	26 (10.8%)
Norway	9 (3.8%)
Sweden	18 (7.5%)
UK	130 (54.2%)
USA	22 (9.2%)
Age at baseline (years)	
Mean	62.3
Median	62.0
Range	27–89
Gender	
Male	141 (58.8%)
Female	99 (41.3%)
Baseline clinical data	
Disease status	
Newly diagnosed	225 (93.8%)
Relapsed/refractory	15 (6.3%)
WHO performance status	
0	87 (36.3%)
1	84 (35.0%)
2	29 (12.0%)
3	23 (9.6%)
4	2 (0.8%)
Missing	15 (6.3%)
β-2 Microglobulin (mg/L)	
<4	86 (35.8%)
⩾4 or <8	67 (27.9%)
⩾8	23 (9.6%)
Missing	64 (26.7%)
Fractures present	
Yes	75 (31.3%)
No	138 (57.5%)
Missing	27 (11.3%)
Treatment received	
Chemotherapy intensity	
Conventional[a]	170 (70.8%)
High[b]	58 (24.2%)
Missing	12 (5.0%)
Transplant received	
Yes	51 (21.3%)
No	166 (69.2%)
Missing	23 (9.6%)
Bisphosphonates taken	
Yes	215 (89.6%)
No	14 (5.8%)
Missing	11 (4.6%)

a Conventional dose chemotherapy includes: MP (melphalan + prednisolone), BP (bendamustin + prednisolone), ID (idarubicin + dexamethasone), CTD (cyclophosphamide + thalidomide + dexamethasone), CVAD (cyclophosphamide + vincristine + doxorubicin (adriamycin) + dexamethasone), lenalidomide + dexamethasone, bortezomib.
b High dose chemotherapy includes regimens given with high dose melphalan: VAD (vincristine + doxorubicin (adriamycin) + dexamethasone), VID (vincristine + idarubicin + dexamethasone), CyDex (cyclophosphamide + dexamethasone), ID (idarubicin + dexamethasone), AD (doxorubicin (adriamycin) + dexamethasone).

of age, gender, disease status, performance status, presence of fractures, β-2 microglobulin, clinical response, chemotherapy dose or receipt of bisphosphonates or transplant. However, they differed significantly by country and trial, probably due to the heterogeneity of trial populations.

Compliance for all QLQ-MY24 items was high, with less than 5% of responses missing. The greatest numbers of missing responses (3–5%) were mainly from items in the side-effects scale at baseline, or the social support scale.

In the debriefing questions, the patients reported a mean completion time of 12 min in total for both questionnaires, with 83% completing within 15 min. 10% of patients at baseline and 18% at follow-up thought there were items not covered by the questionnaires. Most of these were unique to each person although a few reported trembling, eye problems, dizziness and headaches. These issues were considered in the earlier stages of development of the module but were regarded as not being myeloma-specific issues. 5% of patients found some of the questions from the myeloma module confusing or difficult to answer. Comments mainly related to the question 'Were you worried about dying', although most patients had answered this question. A few patients (2.5%) found some of the QLQ-MY24 questions upsetting, these were items 51–54 (body image and future perspective).

3.3. *Distribution of item responses*

For both questionnaires many of the distributions were skewed, particularly where the item questioned a specific symptom or feeling (Table 2). However the full range of responses was seen for all items. The median score for all items in the social support scale was at the ceiling (4) and a large proportion (56.5%) of patients scored the highest possible score for this scale. The side-effects scale also showed an existing level of side-effects at baseline, which was a pre-treatment for most patients.

3.4. *Multi-trait scaling and internal consistency*

Convergent and divergent validity results are shown in Table 3, along with the Cronbach's alpha coefficient to assess the extent to which items in each hypothesised scale are inter-related. Internal consistency was greater than 0.7 in all scales at both time points, suggesting that the reliability is acceptable for all the hypothesised scales.

All scales except side-effects showed good convergence since the correlations between items and their hypothesised scale were greater than 0.4. Scaling successes, where items are correlated significantly higher with their hypothesised scale than another, were seen for all items except

Table 2 – Summary statistics for the myeloma module (QLQ-MY24) items at follow-up

Proposed subscale	Item	Number of responses	Mean	Standard deviation	Median (range)
Disease symptoms	31. Have you had bone aches or pain?	239	2.1	0.9	2 (1–4)
	32. Have you had pain in your back?	239	2.0	0.9	2 (1–4)
	33. Have you had pain in your hip?	240	1.6	0.8	1 (1–4)
	34. Have you had pain in your arm or shoulder?	240	1.5	0.7	1 (1–4)
	35. Have you had pain in your chest?	240	1.3	0.6	1 (1–4)
	36. If you had pain did it increase with activity?	221	1.8	0.9	2 (1–4)
Side-effects of treatment	37. Did you feel drowsy?	237	1.9	0.8	2 (1–4)
	38. Did you feel thirsty?	236	1.8	0.9	2 (1–4)
	39. Have you felt ill?	238	1.8	0.9	2 (1–4)
	40. Have you had a dry mouth?	237	2.0	0.9	2 (1–4)
	41. Have you lost any hair?	237	1.8	1.1	1 (1–4)
	42. Answer this question only if you lost any hair: Were you upset by the loss of your hair?	104	2.0	0.9	2 (1–4)
	43. Did you have tingling hands or feet?	238	1.7	0.8	2 (1–4)
	44. Did you feel restless or agitated?	236	1.7	0.8	2 (1–4)
	45. Have you had acid indigestion or heartburn?	239	1.4	0.7	1 (1–4)
	46. Have you had burning or sore eyes?	238	1.4	0.7	1 (1–4)
Social support	47. Were you satisfied with your relationship with your doctors?	236	3.7	0.6	4 (1–4)
	48. Were you satisfied with the care you received from your doctors?	235	3.7	0.6	4 (1–4)
	49. Were you satisfied with the information you received about your illness?	237	3.5	0.7	4 (1–4)
	50. Did you feel that you were being listened to by your doctor/nurse?	236	3.6	0.7	4 (1–4)
Body image	51. Have you felt physically less attractive as a result of your disease or treatment?	236	2.1	1.0	2 (1–4)
Future perspective	52. Have you been thinking about your illness?	239	2.8	0.8	3 (1–4)
	53. Have you been worried about dying?	240	1.9	0.9	2 (1–4)
	54. Have you worried about your health in the future?	240	2.7	0.9	3 (1–4)

Table 3 – Multi-trait scaling and reliability for the hypothesised scales for the QLQ-MY24

Questionnaire subscale	Baseline results			Follow-up results		
	Item own scale correlation[a]	Item other scale correlation[b]	Cronbach's alpha	Item own scale correlation[a]	Item other scale correlation[b]	Cronbach's alpha
Disease symptoms	0.46–0.77	0.04–0.45	0.76	0.41–0.83	0.01–0.35	0.76
Side-effects of treatment	0.17–0.72	0.01–0.56	0.82	0.33–0.62	0.02–0.49	0.70
Social support	0.71–0.93	0.00–0.27	0.89	0.86–0.91	0.01–0.23	0.92
Body image	–	0.11–0.37	–	–	0.11–0.45	–
Future perspective	0.81–0.88	0.01–0.42	0.79	0.83–0.89	0.07–0.40	0.81

a Corrected for overlap.
b Correlation with other scales from the myeloma module only. The absolute values are displayed.

for some from the side-effects scale. At follow-up (after some treatment) the two items where either criteria was not satisfied were 42 (upset by hair loss) and 45 (acid indigestion) with item-scale correlations of 0.33 and 0.34, respectively. However, item 42 was only applicable to those who had hair loss (item 41) and this had an acceptable item-scale correlation of 0.42. Item 45 was more highly correlated with the side-effects (0.34) scale than with any other scale (0.03–0.11).

3.5. *Relationships between scales from the module and core questionnaires*

Correlations between scales from the QLQ-MY24 and the core questionnaire were examined at baseline and follow-up (Table 4). The correlations between scales are mainly of similar magnitudes at both timepoints. The social support, body image and future perspective scales were weakly correlated with the scales from the QLQ-C30 (correlations less than 0.25, 0.40 and 0.56, respectively), indicating that they are assessing aspects of quality of life in myeloma patients that are not currently covered by the core questionnaire. The disease symptoms scale from the module and the pain scale from the core questionnaire are strongly correlated at both timepoints (0.75–0.78), as was hypothesised at the start of the trial. There are also moderate correlations between the side-effects scale from the module and a number of the QLQ-C30 scales at both time points, in particular all of the functional scales as well as the global health status/QOL and fatigue scales.

4. Resulting changes to the QLQ-MY24

These analyses confirm three of the four hypothesised subscales and the single item. The social support scale was considered not to be contributing to the questionnaire due to the ceiling effect (more than 50% of patients scoring the maximum for the scale). The social support scale scores were not able to significantly discriminate between the sub-groups for either of the clinical parameters (data not shown). Items 47–50 were therefore removed from the questionnaire and reliability and validity of the final QLQ-MY20 questionnaire (Fig. 1) have been demonstrated here.

4.1. *Known-group comparisons*

The QLQ-MY20 scale scores for patients with performance status 0, 1 or 2 (number of patients, $n = 200$) were compared to those with performance status 3 or 4 ($n = 25$) at baseline, and the results can be seen in Table 5. Effect sizes ranged from 0.39 to 1.06. Three of the scales from the module: disease symptoms, side-effects of treatment and body image showed evidence of a significant difference between the patient groups. The future perspective scale showed a trend towards a difference ($p = 0.07$). Note, however, there were very few patients in the poorer performance status group. Approximately two-thirds of the scales from the core questionnaire also showed evidence of significant differences. The mean scores indicate that patients with performance status 0 (19.4) may have different QOL to those with 1, 2, 3 and 4 (38.6, 42.3, 43.6 and 63.9, respectively).

75 patients (35%) were reported as having a fracture at baseline. The disease symptoms and side- effects scales were significantly increased by the presence of fractures at baseline (effect sizes 0.54 and 0.48, respectively), and there was a trend towards a decrease in the future perspective scale (effect size 0.31). Eight of the 15 scales from the core questionnaire also showed a significant difference in baseline quality of life between these patients.

Comparison between newly diagnosed and relapse patients was not possible due to the low number of relapse patients ($n = 15$). No significant differences were found for any scale from either questionnaire for the comparisons between patients with different β-2-microglobulin levels (maximum effect size 0.3). Similarly for the comparison of conventional and high-dose therapy patients, where all effect sizes were less than 0.35, apart from for the constipation scale from the QLQ-C30 which showed significantly worse constipation in the conventional group (effect size 0.67, $p < 0.0001$).

4.2. *Responsiveness to change*

Changes in quality of life over time are expected as a patient's condition changes. This analysis compared the baseline and follow-up scale scores for the 137 (57%) patients who achieved at least a PR (Table 6). Disease symptoms and body image significantly decreased over time in these patients, and the side-

Table 4 – Correlations between scales in the EORTC QLQ-C30 and the QLQ-MY24

	Myeloma module					QLQ-C30														
	DS	SE	SS	BI	FP	PF	RF	EF	CF	SF	QL	FA	NV	PA	DY	SL	AP	CO	DI	FI
Myeloma module																				
DS	–	0.37	0.01	–0.28	–0.24	–0.41	–0.33	–0.29	–0.17	–0.25	–0.38	0.31	0.17	0.75	0.14	0.14	0.15	0.13	0.01	0.13
SE	0.56	–	–0.11	–0.34	–0.39	–0.43	–0.48	–0.56	–0.50	–0.43	–0.49	0.58	0.32	0.30	0.36	0.25	0.35	0.23	0.32	0.25
SS	–0.05	–0.22	–	0.11	0.16	0.03	0.00	0.25	0.13	0.05	0.12	–0.04	0.00	–0.05	–0.11	–0.13	–0.06	0.05	–0.04	0.00
BI	–0.36	–0.37	0.11	–	0.45	0.22	0.19	0.36	0.18	0.28	0.27	–0.17	–0.16	–0.20	–0.05	–0.03	–0.12	–0.03	–0.08	–0.28
FP	–0.27	–0.42	0.05	0.22	–	0.24	0.24	0.55	0.29	0.35	0.32	–0.25	–0.20	–0.18	–0.11	–0.17	–0.14	–0.08	–0.13	–0.31
QLQ-C30																				
PF	–0.57	–0.51	0.11	0.34	0.23	–	0.78	0.36	0.41	0.62	0.63	–0.66	–0.25	–0.50	–0.36	–0.25	–0.39	–0.21	–0.13	–0.26
RF	–0.56	–0.53	0.06	0.31	0.26	0.80	–	0.40	0.42	0.66	0.69	–0.69	–0.30	–0.44	–0.37	–0.22	–0.46	–0.20	–0.18	–0.25
EF	–0.41	–0.64	0.19	0.30	0.56	0.44	0.43	–	0.61	0.47	0.47	–0.49	–0.31	–0.25	–0.28	–0.38	–0.36	–0.19	–0.19	–0.25
CF	–0.47	–0.59	0.19	0.32	0.24	0.43	0.37	0.53	–	0.48	0.43	–0.56	–0.35	–0.13	–0.39	–0.27	–0.36	–0.36	–0.11	–0.22
SF	–0.59	–0.61	0.11	0.36	0.37	0.67	0.77	0.48	0.49	–	0.61	–0.69	–0.29	–0.30	–0.29	–0.21	–0.41	–0.26	–0.23	–0.32
QL	–0.55	–0.61	0.09	0.38	0.39	0.65	0.67	0.54	0.46	0.65	–	–0.65	–0.27	–0.43	–0.30	–0.33	–0.45	–0.14	–0.19	–0.31
FA	0.57	0.70	–0.12	–0.40	–0.30	–0.75	–0.71	–0.55	–0.57	–0.69	–0.71	–	0.33	0.33	0.44	0.30	0.53	0.25	0.24	0.22
NV	0.24	0.43	–0.08	–0.15	–0.19	–0.32	–0.30	–0.31	–0.36	–0.34	–0.36	0.46	–	0.21	0.17	0.29	0.47	0.16	0.32	0.04
PA	0.78	0.55	–0.02	–0.33	–0.26	–0.75	–0.73	–0.42	–0.41	–0.66	–0.64	0.66	0.29	–	0.13	0.19	0.22	0.10	–0.03	0.15
DY	0.25	0.28	0.00	–0.21	–0.15	–0.23	–0.18	–0.16	–0.37	–0.18	–0.23	0.36	0.07	0.19	–	0.23	0.21	0.17	0.19	0.13
SL	0.29	0.32	–0.09	–0.12	–0.24	–0.31	–0.31	–0.30	–0.33	–0.29	–0.32	0.36	0.18	0.33	0.14	–	0.27	0.10	0.20	0.13
AP	0.33	0.50	–0.11	–0.21	–0.22	–0.33	–0.34	–0.39	–0.42	–0.38	–0.37	0.50	0.57	0.36	0.19	0.27	–	0.23	0.22	0.04
CO	0.43	0.35	0.04	–0.23	–0.11	–0.44	–0.43	–0.19	–0.32	–0.43	–0.40	0.42	0.28	0.48	0.14	0.14	0.38	–	–0.11	0.00
DI	0.06	0.09	0.02	–0.07	–0.01	–0.08	–0.09	–0.05	–0.07	–0.11	–0.10	0.13	0.08	0.07	0.12	0.06	0.01	0.13	–	0.08
FI	0.27	0.22	–0.17	–0.13	–0.30	–0.13	–0.22	–0.29	–0.29	–0.25	–0.14	0.20	–0.00	0.22	0.07	0.30	0.09	0.06	0.00	–

Note: Items in the lower left triangle of the diagonal of dashes represent baseline values; items in the upper right triangle of the diagonal of dashes represent follow-up values.
KEY: DS, disease symptoms; SE, side-effects of treatment; SS, social support; BI, body image; FP, future perspective; PF, physical functioning; RF, role functioning; EF, emotional functioning; CF, cognitive functioning; SF, social functioning; QL, Global Health Status/QoL; FA, fatigue; NV, nausea and vomiting; PA, pain; DY, dyspnoea; SL, insomnia; AP, appetite loss; CO, constipation; DI, diarrhoea, FI = financial difficulties.

Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week:		Not at All	A Little	Quite a Bit	Very Much
31.	Have you had bone aches or pain?	1	2	3	4
32.	Have you had pain in your back?	1	2	3	4
33.	Have you had pain in your hip?	1	2	3	4
34.	Have you had pain in your arm or shoulder?	1	2	3	4
35.	Have you had pain in your chest?	1	2	3	4
36.	If you had pain did it increase with activity?	1	2	3	4
37.	Did you feel drowsy?	1	2	3	4
38.	Did you feel thirsty?	1	2	3	4
39.	Have you felt ill?	1	2	3	4
40.	Have you had a dry mouth?	1	2	3	4
41.	Have you lost any hair?	1	2	3	4
42.	Answer this question only if you lost any hair: Were you upset by the loss of your hair?	1	2	3	4
43.	Did you have tingling hands or feet?	1	2	3	4
44.	Did you feel restless or agitated?	1	2	3	4
45.	Have you had acid indigestion or heartburn?	1	2	3	4
46.	Have you had burning or sore eyes?	1	2	3	4
47.	Have you felt physically less attractive as a result of your disease or treatment?	1	2	3	4
48.	Have you been thinking about your illness?	1	2	3	4
49.	Have you been worried about dying?	1	2	3	4
50.	Have you worried about your health in the future?	1	2	3	4

Thank you for completing this questionnaire

Fig. 1 – EORTC myeloma module (QLQ-MY20).

effects of treatment significantly increased. The pain scale was the only QLQ-C30 scale to show a significant change over time in this analysis ($p < 0.001$), although the emotional and social functioning scales showed a trend towards a difference ($p = 0.02$ and 0.01, respectively).

5. Discussion

This study tested the EORTC QLQ-MY24 in an international field-study of patients with myeloma. The study tested the feasibility of prospectively recruiting patients from clinical trials for the validation study. This may be particularly useful for validation of questionnaires in rarer diseases. This approach to validation proved to be feasible and a suitable alternative to a stand-alone validation study.

Analysis demonstrated that this module is reliable. Two of the subscales (disease symptoms, side-effects) and the body image item were able to discriminate between patients with different performance statuses and three were able to discriminate between patients with and without fractures (disease symptoms, side-effects and future perspective). Responsiveness to change was observed in the group of patients achieving at least partial response to the treatment.

The high overall response rate (81%), low missing item rate (5%) and the low number of comments regarding individual items indicates that the questionnaire is well accepted by patients. Most patients completed the QLQ-C30 and MY24 in less than 15 min. Symptoms raised as not covered by the questionnaire were more generally related to the age of the population rather than myeloma specific issues. The study

Table 5 – QLQ-MY20 known-group comparisons at baseline

QLQ-MY20 QoL Scale	Performance status				Fractures			
	0, 1 or 2 (N = 200) Mean (SD)	3 or 4 (N = 25) Mean (SD)	Effect size	Wilcoxon Mann–Whitney p-value	Absent (N = 138) Mean (SD)	Present (N = 75) Mean (SD)	Effect size	Wilcoxon Mann–Whitney p-value
Disease symptoms[a]	30.8 (23.1)	45.2 (22.0)	0.63	0.0029	29.5 (22.4)	41.6 (22.6)	0.54	0.0003
Side-effects of treatment[a]	18.5 (13.9)	33.5 (16.4)	1.06	<.0001	18.3 (15.1)	25.5 (14.9)	0.48	0.0005
Body image[b]	79.8 (29.7)	59.1 (30.7)	0.70	0.0013	76.7 (31.0)	73.3 (31.9)	0.11	0.4452
Future perspective[b]	42.8 (26.0)	32.7 (24.4)	0.39	0.0650	44.4 (27.0)	36.6 (22.8)	0.31	0.0338

N, number of patients.
SD, standard deviation.
a High scores indicate worse symptoms.
b High scores indicate better support/functioning.

Table 6 – QLQ-MY20 responsiveness to change in patients who responded to treatment

QLQ-MY20 QoL Scale	Baseline score (N = 137) Mean (SD)	Follow-up score (N = 137) Mean (SD)	Wilcoxon Mann–Whitney p-value
Disease symptoms[a]	31.9 (23.2)	21.1 (18.7)	0.0001
Side-effects of treatment[a]	20.3 (15.1)	25.5 (15.2)	0.0036
Body image[b]	80.0 (29.8)	63.2 (32.5)	<.0001
Future perspective[b]	42.3 (26.9)	48.3 (25.3)	0.0554

N, number of patients.
SD = standard deviation.
a High scores indicate worse symptoms.
b High scores indicate better support/functioning.

purposefully included patients receiving a wide range of treatment types (bisphosphonates, high-dose therapy and conventional chemotherapy) including some novel agents developed since the questionnaire was designed (thalidomide, bortezomib (Velcade) and lenalidomide (Revlimid)). It is encouraging therefore to note that the MY24 seems to cover issues raised by these newer treatments as well as the more conventional types of treatment.

Limitations of the current study should be acknowledged. Some countries did not have a concurrent trial suitable for inclusion in the validation study, therefore are not represented in this study. A trial was identified for participation in Southern Europe but this failed to recruit. However, the aim was not to evaluate cross-cultural differences but rather to ensure representation from a number of different countries, which was achieved. Demonstrating sensitivity to change was difficult in a sample, where the majority of patients responded to treatment. Analyses showed sensitivity to change in terms of improvement in clinical condition (i.e. responders) but the sample size was too small to investigate deterioration. Longer follow-up may be required to assess responsiveness to change in patients with progression. The ability to discriminate between clinically distinct groups was also hampered by achieving a relatively homogenous sample of patients in the study. There were only 15 patients with relapsed disease therefore comparisons with newly diagnosed patients were not possible. This is representative of clinical trials being conducted at the time of the study which are mainly in newly diagnosed patients, however further validation work in relapse studies that are now running using the QLQ-MY20 or MY24 will be possible. Other known-group comparisons that were considered (β-2-microglobulin levels and intensity of treatment) are less well proven as clinically distinct groups of patients with regards to quality of life. Longer follow-up may also be required to show QOL differences in these patients. Despite low numbers in some subgroups the module is sensitive to change and can discriminate between the well-known clinically different subgroups. The QLQ-MY20 is more responsive to change than the QLQ-C30 in these patients.

A stand-alone validation study may not be possible in rarer diseases such as myeloma. This study successfully recruited the required number of patients in order to test the psychometric properties of the myeloma module utilising existing clinical trials. The study shows that it may be feasible to use this method for validating new QOL questionnaires as an efficient use of patients and data in these diseases or in situations where the majority of patients are entered in clinical trials with QOL assessment anyway. Compliance rates were similar to those found in stand-alone studies. It is, however, important to select trials for inclusion in study such as this carefully. Data quality will be led by the quality of trial management within each trial.

In addition, heterogeneity of patients will be determined by the clinical trials included in the validation study.

There are no other myeloma-specific quality of life questionnaires validated for use in myeloma patients. The suitability of the EORTC QLQ-C30 has been tested in these patients[4,5] and this study has shown that the MY20 is measuring additional aspects of quality of life, particularly issues around body image and future perspective. Further work has been carried out to look at the relationship between the pain subscale of the QLQ-C30 and the disease symptoms scale and shows that the QLQ-MY20 explores different aspects of pain to the QLQ-C30 (data not shown).

In conclusion, QLQ-MY20 is a reliable and valid instrument for measuring quality of life in myeloma patients. It is now recommended that the QLQ-C30 and QLQ-MY20 are used together to measure quality of life in international clinical trials in multiple myeloma. Translations in 18 languages are currently available.

Conflicts of interest statement

None declared.

Source of support

The study was supported by a grant from the EORTC Quality of Life Group. The study was supported in Germany by a grant from Deutsche Krebshilfe, Bonn.

Acknowledgements

We thank all of the patients who participated in this study. We thank Neil Aaronson for review of the manuscript. G.V. is grateful to Cancer Research UK for financial support. We acknowledge the trials providing data for this study: MRC Myeloma IX, UK; CMG 2002, Czech Republic; NMSG 8 and NMSG 12, Norway, Denmark, Sweden; DSMMV and BP vs MP, Germany; 03-041, 03-191 and 03-328, USA. We also thank the National Co-ordinators and the clinical co-ordinators of each participating trial and the staff involved in the provision of data for the study; N Abildgaard, Odense University Hospital, Denmark; S Crofts, Southampton General Hospital, UK; C Fleissner, Universitatsklinikum Charite, Germany; M Hjorth, Sjukhuset i Lidkoping, Sweden; G Juliusson, Lund University Hospital, Sweden; T Kirkpatrick, Dana-Farber Cancer Institute, USA; I Nesthus, Haukeland University Hospital, Norway; O Slama, University Hospital Brno, Czech Republic; A Smith, Southampton General Hospital, UK; A Waage, St Olav's Hospital, Norway; T Young, Mount Vernon Hospital, UK.

REFERENCES

1. Selby P, Gore M. Multiple myeloma. In: Peckham M, Pinedo HM, Veronsei U, editors. *Oxford textbook of oncology*. Oxford: Oxford University Press; 1996. p. 1852–78.
2. Tricot G, Barlogie B, Van Rhee F. Treatment advances in multiple myeloma. *Br J Haematol* 2004;**125**:24–30.
3. Kyle RA, Rajkumar SV. Multiple myeloma. *NE J Med* 2004;**351**:1860–73.
4. Wisloff F, Eika S, Hippe E, et al. Measurement of health-related quality of life in multiple myeloma. Nordic Myeloma Study Group. *Br J Haematol* 1996;**92**:604–13.
5. Wisloff F, Hjorth M. Health-related quality of life assessed before and during chemotherapy predicts for survival in multiple myeloma. Nordic Myeloma Study Group. *Br J Haematol* 1997;**97**:29–37.
6. Sprangers MA, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer. Approach to quality of life assessment: guidelines for developing questionnaire modules. EORTC Study Group on Quality of Life. *Qual Life Res* 1993;**2**:287–95.
7. Cull A, Sprangers MA, Bjordal K, Aaronson NK, on behalf of the EORTC Quality of Life Study Group. EORTC Quality of Life Study Group Translation Procedure. Brussels: EORTC; 1998.
8. Stead ML, Brown JM, Velikova G, et al. Development of an EORTC questionnaire module to be used in health-related quality-of-life assessment for patients with multiple myeloma. European Organization for Research and Treatment of Cancer Study Group on Quality of Life. *Br J Haematol* 1999;**104**:605–11.
9. SAS, Cary, NC: SAS Institute Inc.; 2000.
10. Chronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:296.
11. Nunnally JC. *Psychometric theory*. 2nd ed. New York: McGraw-Hill; 1978.
12. Hays RD, Hayashi T, Carson S, Ware JE. *User's guide for the multitrait analysis of program*. Santa Monica: CA Rand Corporation; 1988.
13. Hays RD, Wang E MultitraitScaling program: MULTI. In: Proceedings of the seventeenth annual SAS Users Group International Co.