Health related quality of life assessment instruments: A prospective study on preference and acceptability among cancer patients referred for radiotherapy

Madelon C.G. Pijls-Johannesma a,*, Anouk Pijpe b, Gertrudis I.J.M. Kempen d, Philippe Lambin a,c, Pieter C. Dagnelie b

a Department of Radiation Oncology (MAASTRO clinic), GROW, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands
b Department of Epidemiology, Faculty of Health Sciences, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
c University Hospital Maastricht, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands
d Department of Health Care Studies, Faculty of Health Sciences, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Received 7 April 2005; received in revised form 2 June 2005; accepted 15 June 2005
Available online 6 September 2005

Abstract

The aim of this study was to determine the preferences for and acceptability of existing Quality of Life (QoL) questionnaires among cancer patients. Sixty-two patients completed a set of questionnaires. Criteria were the percentage of missing values, difficult items and the preferences for the answer options.

Results indicated small proportions of missing values and difficult items for the EORTC QLQ-C30, the MFI, RSCL, and the MQ (0.0–1.0%). Considerably higher proportions of missing values and difficult items were found for the VAS and the PFS (2.5–3.4%). The answer option of the RSCL subscale ‘activity level’ was most preferred (29% of subjects), whereas the VAS and PFS were least preferred (19% and 21%, respectively).

Our results indicate that patients prefer answer categories that have descriptors in words or a narrow range of answer options, rather than non-described answer options or a broad range of answer options per question.

Keywords: Quality of Life; Assessment; Preference; Acceptability; Cancer; VAS; MFI; EORTC QLQ-C30; RSCL; PFS

1. Introduction

Health related quality of life (HRQoL) can be defined as the level of function and well-being with respect to the physical, psychological and social domain [1,2]. HRQoL assessment is not only used to evaluate the effect of clinical and psychosocial interventions, but also as a source of information about the range of patients’ physical, functional and psychosocial problems that may contribute to an improved physician-patient communication [3]. Especially in cancer, a disease that has a major impact on patients’ lives, the importance of HRQoL assessment is generally accepted.

The usual method of collecting HRQoL information is by self-assessment questionnaires. During the last two decades, a number of these questionnaires have been developed [4–8]. So far, evaluations have mainly focused on psychometric properties of such HRQoL scales and not on the patient’s preferences and acceptability of these scales. Preference and acceptability are nevertheless important issues for several reasons. First, the
assessment of HRQoL in cancer patients is often hampered by the incompletion of scales, a phenomenon that can be caused by questions that are unclear to patients or questions that are considered as offensive to patients. Second, if these instruments are not filled out in the correct manner, this may lead to incorrect conclusions. Third, in the setting of a busy outpatient clinic with limited staff, HRQoL instruments must be self-instructive, easily understandable and reasonably convenient for the patient in terms of time and effort. Recently, a study has been published on patient preferences for and acceptability of a number of popular validated self-assessment HRQoL scales amongst breast and lung cancer patients referred for radiotherapy with curative intention in an outpatient oncology clinic.

2. Patients and methods

2.1. Study population

Eligible patients were those with any stage of breast cancer or lung cancer who had been selected for high-dose radiotherapy (≥50 Gy). Patients, having undergone surgery for lung cancer were excluded. From June 2002 to October 2002, we invited 70 patients, older than 18 years of age and with World Health Organization (WHO) performance status 0–2 to participate [10]. Prior to radiotherapy, patients were asked to complete a set of questionnaires that included a demographic profile and several existing multidimensional validated questionnaires (see below). After completion of the set of questionnaires, patients were asked to fill out an evaluation form. The study was approved by the local ethics committee by complying to Dutch legislation and all patients gave written informed consent before the start of the study.

2.2. Measurement tool

The set of questionnaires or scales (which we will call below ‘The measurement tool’), was composed of a Visual Analogue Scale. Five already existing multidimensional validated scales, all of which had previously been used in clinical practice to measure HRQoL and more specifically fatigue as it is one of the most important aspects of HRQoL in cancer patients [11–13]. Therefore, we chose to include especially those (sub-)scales in which the concept of fatigue is represented. Since a recent study showed no order effects on the assessment of HRQoL in cancer patients [14], and more recently a study showed that the order of administration should be consistent throughout the study [15], all patients received the identical measurement tool, as presented in Table 1.

The Visual Analogue Scale (VAS) consists of a 10 cm straight line without scale markers and with descriptors at each end. One VAS item on the degree of fatigue (Table 1) was included in the measurement tool.

The Multidimensional Fatigue Inventory (MFI) is a 20-item scale designed to measure multiple aspects of fatigue experienced during the last few days [16]. The MFI is based on the notion of fatigue as a multidimensional experience, which can be expressed through five subscales of four items each. English and Dutch versions of the MFI have previously been validated in cancer patients undergoing radiotherapy [16–18]. All subscales of the MFI were included in the measurement tool.

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30), version 3.0, is a 30-item cancer-specific scale addressing various aspects of HRQoL [19]. It comprises of five functioning subscales (physical, role, emotional, cognitive and social functioning), three multi-item symptom scales (fatigue, pain and nausea/vomiting), two global health/quality of life subscales, and a number of single items addressing various symptoms and perceived financial impact. All subscales were included in the measurement tool.

The Rotterdam Symptom Checklist (RSCL) is a 39-item scale, which assesses symptoms during the preceding week [20]. The reliability and validity of the RSCL in cancer patients have been confirmed previously [21] [7,22]. The RSCL assesses four domains: physical symptoms, psychological distress, activity level, and overall quality of life. Because of major overlap with items in other questionnaires included in the measurement tool, the subscale ‘physical symptoms’ was excluded.

The Maastricht Questionnaire (MQ) was developed at Maastricht University [23] to measure vital exhaustion. The scale is one-dimensional and comprises of 23 items on abnormal fatigue, loss of energy, increased irritability, and feelings of demoralization. All items were included in the measurement tool.

The Piper Fatigue Scale (PFS) [24] is composed of 22 numerical items which assess fatigue as experienced at present. The PFS measures four dimensions of subjective fatigue: ‘behavioural/severity’, relating to the severity, distress, and degree of disruption in activity of daily living; ‘affective meaning’, relating to the emotional meaning attributed to fatigue; ‘sensory’, relating to the physical symptoms of fatigue; and ‘cognitive/mood’, relating to mental and mood states. All items
were included in the measurement tool. For the present study, the PFS was translated into Dutch independently by three native speaking Dutch researchers. Thereafter, the translations of the three researchers were discussed in several consensus meetings that included all authors except AP and PL.

Table 1
Overview of used scales: subscales, examples of items and answer options

<table>
<thead>
<tr>
<th>Scales*</th>
<th>Subscale(s) (# of items)</th>
<th>Example (arbitrary) of item and answer options</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS</td>
<td>Fatigue (1)</td>
<td>To what extent did you feel tired during the past week?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 not tired at all</td>
</tr>
<tr>
<td>MFI</td>
<td>- General fatigue (4)</td>
<td>I feel fit (item 1) yes, that is true no, that is not true</td>
</tr>
<tr>
<td></td>
<td>- Physical fatigue (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Reduced activity (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Reduced motivation (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Mental fatigue (4)</td>
<td></td>
</tr>
<tr>
<td>EORTC-</td>
<td>Symptom and Functional</td>
<td>Did you feel sick?</td>
</tr>
<tr>
<td>QLQ-C30</td>
<td>scales (28)</td>
<td>(item 14)</td>
</tr>
<tr>
<td></td>
<td>- Global health/ QoL (2)</td>
<td>How would you rate your overall quality of life during the past week? (item 30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>very poor</td>
</tr>
<tr>
<td>RSCL</td>
<td>Psychological distress (7)</td>
<td>Have you, during the past week, been bothered by:</td>
</tr>
<tr>
<td></td>
<td>Activity level (8)</td>
<td>Depressed mood (item 3) not at all a little quite a bit very much</td>
</tr>
<tr>
<td></td>
<td>Overall QoL (1)</td>
<td>Would you please mark the answer that applies most to your condition of the past week:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Go to work (item 8) unable only with help without help without help with difficulty</td>
</tr>
<tr>
<td>MQ</td>
<td>Vital exhaustion (23)</td>
<td>All things considered, how would you describe your quality of life during the past week?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>excellent good moderately good neither good nor bad rather poor poor extremely poor</td>
</tr>
<tr>
<td>PFS</td>
<td>- Behavioral/severity (6)</td>
<td>To what degree are you now feeling (item 13)</td>
</tr>
<tr>
<td></td>
<td>- Affective meaning (5)</td>
<td>awake sleepy</td>
</tr>
<tr>
<td></td>
<td>- Sensory (5)</td>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td></td>
<td>- Cognitive/mood (6)</td>
<td></td>
</tr>
</tbody>
</table>

* Abbreviations explained in Section 2.2.
2.3. Preferences and acceptability

To evaluate the preferences and acceptability of the measurement tool, three approaches were followed. First, patients were asked to mark, during completion of the measurement tool, the items they found difficult to answer. Second, we analysed the numbers of missing values for each item, assuming that a high proportion of missing values partly reflects a low understanding or a low acceptability of that specific item or (sub-)scale. Third, patients were asked, after completion of the measurement tool, to fill out an evaluation form on which they could indicate which answer options they preferred the least and the most. Next to it, the number of patients that needed assistance (read questions aloud) in completing the measurement tool was registered.

2.4. Outcome measures and statistical analysis

Student t-test and Fisher’s Exact test were used to compare differences in continuous and categorical variables, respectively, between tumour groups. Evaluation of the measurement tool was first done by calculating the proportion of missing values per item and per (sub-)scale:

\[
\text{Proportion of missing values} = \frac{\text{Number of missing values}}{\text{Total number of items}} \times 100.
\]

Second, the proportion of items marked as difficult to answer per (sub-)scale was calculated:

\[
\text{Proportion of difficult items} = \frac{\text{Number of difficult items}}{\text{Total number of items}} \times 100.
\]

Third, Pearson’s correlations were used to evaluate the association between missing values and items rated as difficult. Finally, the proportion of patients that had indicated a specific answer option as the most and the least preferred was computed. The level of statistical significance was set at 0.05 (two-sided).

3. Results

3.1. Patient characteristics

Out of 70 patients invited for the study, 62 (89%) completed the measurement tool. Eight patients, of whom seven were lung cancer patients, refused to cooperate because they were too tired. The characteristics of the 62 patients who completed the study are presented in Table 2. Lung cancer patients, compared to breast cancer patients, were older \(P = 0.014\), had a lower Body Mass Index \(P = 0.005\) and a poorer WHO performance status \(P = 0.014\). From the 27 lung cancer patients, 23 had non-small-cell lung cancer (stage I: \(n = 4\), stage II: \(n = 5\), stage III: \(n = 10\), and not staged: \(n = 2\)), according to the TNM staging classification system [25] (results not tabulated), and six patients had small-cell lung cancer (limited disease). Of the 35 breast cancer patients, 23 patients had stage I, 9 patients stage II, 1 patient stage III, and 2 patients were not staged (TNM, not tabulated). Compared to patients with breast cancer, relatively more lung cancer patients had been pre-treated with chemotherapy \(P = 0.021\).

3.2. Preference and acceptability

In Table 3, the number of patients with a certain range of missing values per (sub-)scale, and the proportion of missing values, are presented per (sub-)scale. The (sub-)scales with the smallest proportion of missing values (<1%) were the MFI, the EORTC QLQ-C30, the subscales ‘psychological distress and overall QoL’ of the RSCL, and the MQ. A relatively large proportion of missing values was observed in the VAS (3.2%), the RSCL subscale ‘activity level’ (4.2%) and the PFS (2.5%). The high proportion for the PFS was attributed to the subscales ‘behavioural/severity’ (3.1%) and ‘affective meaning’ (6.6%) (not tabulated). Items with missing values in five or more patients were item 8 of the RSCL subscale ‘activity level’ \(n = 20\), ability to go to work), item 9 of the MQ \(n = 6\), sexual activities), and the PFS-items 4 \(n = 7\), sexual activities), item 9 \(n = 5\), evaluation of fatigue on a 10-point scale with descriptors “protective-destructive”) and item 10 \(n = 5\), evaluation of fatigue on a 10-point scale with descriptors “positive-negative”).
Also presented in Table 3, is the proportion of items indicated by patients as ‘difficult to answer’. (Sub-)scales with a relatively small proportion of items marked as difficult (<1%) were the MFI, the EORTC QLQ-C30 subscales ‘symptom and functioning’, and the RSCL subscales. In contrast, patients’ evaluation showed a relatively large proportion of difficult items for the subscale ‘global health’ of the EORTC QLQ-C30 (6.3%), and for the PFS (3.4%). Also a relatively large proportion of the patients (3.1%) evaluated the VAS as difficult.

The MFI and two subscales of the RSCL (‘activity level’, ‘distress’) had hardly any items marked by patients as difficult. In contrast, a considerable number of items were marked as difficult to answer, for the EORTC QLQ-C30, the MQ and the PFS. Within the EORTC QLQ-C30, patients had most difficulties with the two items (both rated on a 7-point scale) of the subscale ‘global health/QoL’, which refer to the patients’ opinion on their overall well-being and quality of life. For the MQ, items marked as difficult to answer were scattered across the scale, except for item 9 (referring to whether patients enjoy sexual activities nowadays as much as in the past), which was marked as difficult by seven patients. Of the PFS, items marked most frequently as difficult to answer were items 4 (n = 8) and 9 (n = 8). Item 4 refers to the interference of fatigue with sexual activity (10-point scale with descriptors “not at all – very strong”) and item 9 refers to the evaluation of fatigue (10-point scale with descriptors “protective – destructive”).

Overall, a significant correlation (r = 0.27, P = 0.004) was found between the proportion of missing values per item and the number of times this item was evaluated as difficult by patients (Table 3). However, this was not necessarily the case for all individual items. Separately from the total correlation, the correlation of the three items with the highest number of missing values and number of times these items were evaluated as difficult, were calculated.

### 4. Discussion

To our knowledge, this is the first report on preference for and acceptability of self-assessment health related quality of life questionnaires among cancer patients prior to curative radiotherapy: The following tentative conclusions can be drawn: (1) patients appear to prefer answer options that have been described in words (e.g. RSCL) or have no label at all (e.g. the PFS) and (2) patients appear to prefer questions with a narrow range of answer options (maximum of 4 or 5), and to dislike broad range answer formats (e.g. 10 points or VAS). In addition, if the topic of certain items is either not applicable (for example the assessment of the ability to go to work when a patient already has retired) or too sensitive (e.g. sexual activities), this may lead to a relatively large number of missing answers.

Despite the unique approach in the present study, several limitations must be noted. First, since the study population is relatively small and the number of questionnaires is substantial which may have affected the reliability of our results, our findings should be confirmed in subsequent larger studies. Second, as we included lung and breast cancer patients selected for curative radiotherapy, it could be argued that our conclusions may only apply to these specific patient groups. However, despite differences between lung and breast cancer patients with respect to age, BMI, WHO performance status, tumour stage and pre-treatment with chemotherapy, the overall results with regard to preference and acceptability were similar in both tumour groups, which would suggest that different patient groups have similar feelings about these HRQoL assessment tools. Third, the number of HRQoL assessment tools that are available to cancer patients is numerous. For reasons of subject burden, we could include only a relatively small selection of these in the present study. We chose a number of well-known and well-validated

### Table 3

<table>
<thead>
<tr>
<th>(Sub)scale</th>
<th>Number of items</th>
<th>Number of patients categorized by total number of missing values (0–4)</th>
<th>Missing values (%)</th>
<th>Difficult items (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS</td>
<td>1</td>
<td>60 2 0 0 0</td>
<td>3.2</td>
<td>3.1</td>
</tr>
<tr>
<td>MFI</td>
<td>20</td>
<td>60 1 1 0 0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>EORTC-QLQ-symptom and functioning</td>
<td>28</td>
<td>56 5 1 0 0</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>EORTC-QLQ-global health</td>
<td>2</td>
<td>60 1 0 0 0</td>
<td>0.8</td>
<td>6.3</td>
</tr>
<tr>
<td>RSCL-psychological distress</td>
<td>7</td>
<td>61 0 0 1 0</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>RSCL-activity level</td>
<td>8</td>
<td>42 19 1 0 0</td>
<td>4.2</td>
<td>0.6</td>
</tr>
<tr>
<td>RSCL-overall QoL</td>
<td>1</td>
<td>62 0 0 0 0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>MQ</td>
<td>23</td>
<td>54 6 2 0 0</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>PFS</td>
<td>22</td>
<td>47 5 5 1 4</td>
<td>2.5</td>
<td>3.4</td>
</tr>
</tbody>
</table>

* Formulas used to calculate missing values and difficult items are described in the section Outcome measures and statistical analysis.
(sub-)scales, covering the most frequently reported symptoms of HRQoL, and its sub-domain fatigue, in cancer patients.

There may be several reasons why patients do not fill out a question in self-assessment HRQoL questionnaires. The topic of the question may not be applicable (anymore) to the patient, the answer option may be found difficult, the patient may find the subject of the question hard to empathize with, or he/she may consider it offensive or too sensitive. Item 8 of the RSCL-'activity level', was missed 20 times and was therefore responsible for the highest proportion of missing values (4.2%) of this subscale within the measurement tool. An explanation for this result is that this item refers to the ability to go to work, whereas the majority of patients in the present study were either retired or had no paid job. Other examples of a high number of missing values due to the subject of the specific question were questions related to sexual activities (item 9 of the MQ and item 4 of the PFS). This subject may be too sensitive for an answer for a considerable proportion of the present study population. Furthermore, the relatively large number of missing values of the PFS subscale ‘affective meaning’ (6.6%, not tabulated) is remarkable. Items in this subscale refer to the extent to which fatigue is experienced by patients as agreeable vs. disagreeable, protective vs. destructive, positive vs. negative, and normal vs. abnormal. Apparently, these items are not offensive, but the way the questions are formulated or the concept covered by the four items of this subscale may very well be difficult to grasp for a considerable proportion of patients.

In the present study, a pattern of the type of answer options with regard to the proportion of missing and difficult items was identified. The number of missing values in scales with a “broad range” of answer categories per item (range: 0.8–3.1%) was relatively high when compared to the number of missing values in scales with a “narrow range” of answer categories (range 0.0–0.8%). This pattern also applies to the proportion of difficult items per scale, with exception of the RSCL-'overall QoL’ (0.0%). All scales with a “broad range” of answer options appeared to have a considerably higher proportion of items indicated as difficult to answer (range: 3.1–6.3%), when compared to scales with a “narrow range” of answer options per item (range: 0.0–1.0%). A plausible explanation for the observation that the subscale ‘overall QoL’ of the RSCL is an exception (one item, never missed), is that all answer options of this 7-point sub-scale are described in words in the scale. This would suggest that, when answer options are described in words, the number of answer categories may be of lesser importance. However, it must be emphasized that a considerable part of the patients indicated that the kind of answer option made no difference to them.

The total time needed to fill out the measurement tool did not differ between the tumour groups. However, whereas none of the breast cancer patients needed assistance in completing the measurement tool, five lung cancer patients, all with performance status WHO-2, needed assistance to complete the measurement tool. Although this finding is based on a small number of patients, the possibility must be taken into account that a relatively large proportion of patients with poor performance status (≥2) may need assistance in order to complete HRQoL assessment.

In conclusion, the results of the present study indicate that, when evaluating the effect of clinical and psychosocial interventions in cancer patients, respondents may prefer HRQoL measurement tools with answer categories that have either descriptors in words, or a relatively narrow range of answer options per question. Furthermore, to obtain valid data from HRQoL scales, the characteristics of the study population must be taken into account, in order to estimate if questions are either not applicable, too sensitive, or offensive.

These findings may help in designing and choosing patient-friendly HRQoL scales, lead to a reduction in non-response or missing items in HRQoL assessment.

Conflict of interest statement

None declared.

References


