Listen to their answers! Response behaviour in the measurement of physical and role functioning

**Marjan J. Westerman · Tony Hak · Mirjam A. G. Sprangers ·
Harry J. M. Groen · Gerrit van der Wal · Anne-Mei The**

Abstract

Background Quality of life (QoL) is considered to be an indispensable outcome measure of curative and palliative treatment. However, QoL research often yields findings that raise questions about what QoL measurement instruments actually assess and how the scores should be interpreted.

Objective To investigate how patients interpret and respond to questions on the EORTC-QLQ-C30 over time and to find explanations to account for counterintuitive findings in QoL measurement.

Methods Qualitative investigation was made of the response behaviour of small-cell lung cancer patients ($n = 23$) in the measurement of QoL with the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30). Focus was on physical functioning (PF, items 1 to 5), role functioning (RF, items 6 and 7), global health and QoL rating (GH/QOL, items 29 and 30). Interviews were held at four points: at the start of the chemotherapy, 4 weeks later, at the end, and 6 weeks after the end of chemotherapy. Patients were asked to 'think aloud' when filling in the questionnaire.

Results Patients used various response strategies when answering questions about problems and limitations in functioning, which impacted the accuracy of the scale. Patients had scores suggesting they were less limited than they actually were by taking the wording of questions literally, by guessing their functioning in activities that they did not perform, and by ignoring or excluding certain activities that they could not perform.

Conclusion Terminally ill patients evaluate their functioning in terms of what they perceive to be normal under the circumstances. Their answers can be interpreted in terms of change in the appraisal process (Rapkin and Schwartz 2004; *Health and Quality of Life Outcomes, 2*, 14). More care should be taken in assessing the quality of a set of questions about physical and role functioning.

M. J. Westerman · G. van der Wal · A.-M. The
Department of Public and Occupational Health, EMGO Institute,
VU University Medical Centre, Amsterdam, The Netherlands

M. J. Westerman (✉)
Department of Methodology and Statistics, Institute of Health
Sciences, VU University, De Boelelaan 1085, 1081 HV
Amsterdam, The Netherlands
e-mail: marjan.westerman@falw.vu.nl

T. Hak
RSM, Erasmus University, Rotterdam, The Netherlands

M. A. G. Sprangers
Department of Medical Psychology, Academic Medical Centre,
Amsterdam, The Netherlands

H. J. M. Groen
Department of Pulmonary Diseases, University Hospital,
Groningen, The Netherlands

Introduction

Health-related quality of life (QoL) is considered to be an important outcome for evaluating the impact of disease and for assessing the effectiveness of treatment. QoL is particularly relevant for cancer patients who are willing to undergo risky and toxic treatment [1–4]. This is especially so in patients with small-cell lung cancer (SCLC), for whom therapy is unlikely to be curative and who receive life-prolonging and 'palliative' therapy. Understanding the burden of symptoms and the relative effects of chemotherapy on a

person's QoL is important for decision-making and in daily clinical practice for optimising the QoL of these patients throughout the course of their illness [5].

However, QoL research often yields findings that raise questions about what QOL measurement instruments actually assess and how the scores should be interpreted. For example, Groen et al. studied patients with inoperable non-SCLC treated by radiation with and without chemotherapy [6]. QoL was measured with the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30) [7], which has been designed specifically for use in clinical trials focusing on cancer patients. Contrary to expectations, they found no significant deterioration in the scores over the 6-week treatment period. Patients with a life-threatening disease reported a stable QoL, and the level of QoL of patients with a severe chronic illness was found to be no worse or no better than that of less severely ill patients or healthy people [8–14]. Furthermore, several studies show that patients' own evaluations may differ considerably from those made by clinicians and significant others [15, 16]. Such findings, labelled by Breetvelt and Van Dam as 'under-reporting of problems', suggest that patients report less distress and dissatisfaction than they actually experience [17]. In short, QOL measures do not consistently distinguish illness experiences, they show little convergence across measurement perspectives, and they are often only weakly related to objective criteria, i.e., negative factors such as side effects of chemotherapy have little effect on the patients's QoL (the so-called satisfaction paradox) [18, 19].

In recent years the response shift theory has gained increasing attention [20]. Response shift refers to a change in a patient's internal standards, values, and conceptualization of QoL. It is suggested that patients make the best of their condition by coping, rethinking, and reframing their experiences, and that this adaptive self-regulation may explain these discrepancies [21–23].

In order to examine response shift more closely, we used the EORTC-QLQ-C30 and the SEIQoL-DW (a patient-centred measure to assess individual QoL) in a qualitative exploratory longitudinal study involving 31 patients. Recently, we reported our results on the measurement of response shift with SEIQoL-DW (i.e., change in values and conceptualization) [24, 25] and also our findings about how patients dealt with the EORTC question 'were you tired' [26]. We found discrepancies between the levels of fatigue as measured with the questionnaire and the levels spontaneously reported during the interview. These findings seem to be in line with a study of Cox [27] who had studied cancer patients who participated in a clinical trial; she found no significant change in the EORTC assessment, but the in-depth interviews led to alternative conclusions about

the impact of the trial on these patients. What we learned both from the literature and our own work is that respondents are engaged in complex response patterns when filling out a questionnaire. We argued that it is important to know more about 'what actually happens' in QoL measurement before trying to measure any response shift. This paper describes the results of our longitudinal multiple case study in which we investigated how patients interpret and respond to questions on the EORTC-QLQ-C30 over time and in which we focus on global health, global QoL (GH/QOL), and items on the physical (PF) and role (RF) functioning scales. This paper is a follow-up to this study with the aim to search for explanations to account for the above-mentioned 'poorly understood' QoL outcomes.

Methods

Procedures and study sample

In 2000, approval was obtained from the Medical Ethics Committee of the VU University Medical Center for our study, "Response shift in quality of life in the palliative treatment of small-cell lung cancer patients". In March 2001 we started to recruit SCLC patients in five outpatient clinics for chest diseases in the Netherlands. All patients were evaluated for first line chemotherapy, and no restrictions were made with regard to age or treatment (chemotherapy or a combination of chemotherapy and radiotherapy). The participating patients gave written informed consent and were interviewed during the treatment trajectory. The first interview was carried out within 7–10 days after diagnosis at the start of the chemotherapy. In the original study design, the second interview was planned after completion of the course of chemotherapy. However, after the inclusion and first interviews with three patients, we decided that we would also interview the patients during the treatment. Therefore, the second interview was held 4 weeks after the first interview, the third 7–10 days after completion of the chemotherapy treatment, and the fourth interview 6 weeks after completion of treatment.

Between March 2001 and September 2003, 41 eligible patients were invited to participate in the study. Four patients were unwilling to participate, and six were not interviewed because of their imminent death. Of the 31 patients who were interviewed, eight were excluded from the analysis because their data were incomplete, i.e., they were only interviewed once (six died within a month after the first interview and two were too sick at the second and died before the end of the planned chemotherapy). Consequently, the final study sample consisted of 23 SCLC patients, 12 of whom were diagnosed with limited disease

(3 male and 9 female, mean age 55, range 42–69) and 11 with extensive disease (8 male and 3 female, mean age 64, range 39–72). Except for seven patients whose chemotherapy was combined with local radiation of the tumour, all patients received standard chemotherapy. The majority of the patients were married (19, 83%) and had children (17, 74%).

Of the 23 patients in our study sample, 15 were interviewed four times, 7 were interviewed three times, and 1 patient was only interviewed twice, resulting in a total of 83 interviews. The interviews were conducted by MW with a duration of 80–110 min. Except for the first interview with three of the patients, all interviews were held in the patient's home.

Materials and qualitative method

In this exploratory, longitudinal multiple-case study, QoL was assessed with the EORTC QLQ-C30 (version 3.0) [7] and the lung cancer module QLQ-CL13 [28]. The EORTC QLQ-C30 is the cancer-specific QoL measurement instrument that is most widely used in European clinical trials. The questionnaire consists of several functioning scales that measure, among other things, physical and role functioning, mental and general health, and global QoL. Furthermore, it measures different symptoms such as pain, dyspnoea, nausea, and fatigue. For the questions about functioning the respondent has four response options. The respondent circles the most appropriate number, i.e., not at all [1], a little [2], quite a bit [3], and very much [4]. General health and global QoL are rated by circling a number between 1 and 7 (respectively, from very poor to excellent). The QLQ-CL13 module measures the extent to which patients experience symptoms or problems related to their lung cancer, such as coughing and hair loss.

The EORTC QLQ-C30 and the CL13 were completed in combination with the Three-Step Test Interview (TSTI) to investigate how respondents interpreted the items and how they responded to them [29]. The TSTI consists of the following steps: (1) concurrent thinking aloud, aimed at collecting observational data on how respondents complete the questionnaire, expressing their thoughts aloud; (2) focussed interview, aimed at clarifying the respondent's previous expression of thoughts while completing the questionnaire; (3) semi-structured interview, aimed at eliciting the respondent's experiences and opinions with regard to the questionnaire.

Interview protocol

In the first interview, QoL was measured with the EORTC QLQ-C30, followed by the lung cancer module QLQ-CL13.

The questionnaires were completed in a concurrent think aloud manner, and after completion the respondents were asked to clarify previous hesitations, and expressions when rating certain items and experiences (i.e., second and third step of the TSTI). Individual QoL was then measured with the Schedule for the Evaluation of Individual Quality of Life–Direct Weighting (SEIQoL-DW) [30, 31]. Finally, we encouraged patients to talk freely about their experiences with the treatment. At follow-up (i.e., at the second, third and fourth interviews) the protocol of the first interview was extended with an extra QoL measurement, i.e., after the SEIQoL-DW assessment, EORTC questionnaires were completed again as a so-called 'then test' [32–34] (i.e., the patients filled in these questionnaires according to how they perceived themselves at the time of the previous interview). In the second and following interviews, EORTC questionnaires (contemporary and then-test QoL assessment) were also completed in a concurrent think-aloud manner. However, in contrast to the first interview, we integrated the second and third step of the TSTI in the assessments. In actual practice this means that we encouraged patients to think aloud, and we probed for clarification after each item if extra information was considered to be useful in order to understand the patient's answers. We adopted a flexible approach in order not to interrupt the natural flow of both the assessment and the patient-interviewer communication.

Interviews were audio-taped and transcribed verbatim. In this article we focus on the physical functioning scale (PF, items 1 to 5), the role functioning scale (RF, items 6 and 7), and global health and QoL rating (GH/QOL, items 29 and 30).

Analysis

Four types of data were collected and used for the analysis [35]: (1) completed questionnaires of all interviews, (2) observed patient behaviour recorded in field notes, (3) transcriptions of the interviews, and (4) memos with background information regarding the illness trajectory of each patient. In order to manage the data of the interviews ($n = 83$), we used the qualitative computer package Kwalitan 5.0 (http://www.kwalitan.net) to extract relevant parts of the transcriptions: (1) think aloud of the GH/QOL questions and the functioning scales (i.e., physical and role) and (2) 'comments' related to the way in which patients were functioning, and the perceived impact of treatment on QoL. Furthermore, in order to deal with the still remaining large amount of extracted data, two of the authors (MW, AT) condensed extracts from transcripts of the 'comments' into core texts. The data (e.g., think aloud combined with scores and core texts of comments) were organised and analysed for each patient separately, resulting in 23 case studies [36]. Initially, three authors (MW,

AT, TH) independently analysed the data of the first two interviews of the first patient. The aim was to understand why a specific response category was chosen and to identify change in the response behaviour by comparing the results of the two interviews. Further analyses of the 23 case studies were conducted by MW. She explored how the patient had answered each item (i.e., interpretation of the question, response, and choice of response category). She interpreted the individual score and the think-aloud data of that item against the background of the patient's illness trajectory and assessed whether that item had received a score that one would expect if it truly reflected the limitations the patient experienced. Finally she looked for change in the process of appraising over time. Results were discussed with AT and, after completing the individual cases, MW and AT searched for patterns in the response strategies for each item by examining similarities and differences among the cases. The robustness of their interpretations was critically discussed by the research team (MW, AT, TH, MS).

Results

Quality of Life during first line chemotherapy

For all patients, chemotherapy consisted of five cycles. After the first cycle (second interview, T2), 13 patients said that they were doing well, considering the circumstances. Patients diagnosed with extensive disease who suffered from tumour-related symptoms, such as dyspnoea and coughing, reported feeling much better than before the treatment. Furthermore, three of the seven patients who had been treated with chemotherapy and radiation therapy experienced a severe physical burden of the side effects of radiation therapy (e.g., pain, problems with eating and drinking, and consequent loss of weight). All the patients experienced the impact of every new cycle as more and more severe. Six weeks after the last cycle, three patients had died, and two were confronted with a recurrence of the cancer and further treatment. The others attempted to pick up their normal life again.

According to the results of the EORTC-QLQ-C30, QoL was affected very little by the chemotherapy (Table 1). The mean values for GH/QOL at the start (first interview, T1) and at the end of the chemotherapy (third interview, T3) were both 66. The mean value of the physical functioning scale decreased from 75 at T1 to 72 at T3, but the mean value of role functioning increased from 62 at T1 to 72 at T3; seven patients had the highest possible rating (100), and only one patient had the lowest possible rating (0).

We were not able to interview all 23 patients four times, and therefore the mean values presented for the second and fourth interview (i.e., T2 and T4) cannot be adequately

compared with the mean values for the first and third interview (i.e., T1 and T3). However, despite these differences and the large individual variations, most of the patients perceived their QoL to be at a higher level at the second interview than at the first (i.e., at the start of chemotherapy): 12 of 19 patients reported a higher level GH/QoL, 2 an equal level, and 5 a lower level of GH/QOL. Patients also reported better QoL at 6 weeks after the course (T4) than at the end of the course (T3): 10 of 18 patients reported a higher level GH/QoL, 4 an equal level, and 4 a lower level of GH/QOL. With respect to the functional scales, the mean values of physical and role functioning also increased 6 weeks after the course, suggesting that most patients were not impaired by the treatment: 8 of the 18 patients reported the highest level (100) of role functioning and only 1 patient the lowest possible rating (0). In order to understand how QoL outcomes in our study should be interpreted, background knowledge of the response behaviour described below has to be taken into consideration.

Patients' answers at the first QoL assessment

Questions 1 to 7 on the QoL instrument reveals important clues to how patients report physical and role limitations. Patients responded in unexpected ways: by focusing on one aspect of the question, by taking the wording of the question literally, and by ignoring or excluding certain activities that they could not perform. This resulted in patients producing QoL scores that suggested they were less limited than they actually were (see Fig. 1 for examples). Furthermore, a few patients guessed their level of functioning in activities that they did not perform or used the strategy *"I didn't do it, so I don't have any trouble"* (example see item 2), and a few compared present with previous experiences (e.g., *"I feel not too bad, compared to last week"*) or with expectations (e.g., *"I feel better as expected"*). The following examples illustrate for each question how, during the first interview (T1), a certain strategy resulted in a different QoL score than would be expected.

Item 1: Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or suitcase? A patient who could hardly walk outside the house answered this question with 'not at all' by arguing that you can't have any trouble if you don't do any shopping: *"I never carry a shopping bag, my wife does the shopping"* [M, age 72, T1]. At T1, 17 out of 23 patients interpreted this item more or less literally and focused on just shopping bags or suitcases: *"I can lift the shopping bag but I can't walk with it"* or *"my suitcase has wheels, so I don't have to carry it"*. Out of the other six patients, four were thinking about other strenuous activities when evaluating this item, and two were just circling the number that corresponded with

Table 1 Individual and mean scores of patients answering the EORTC QLQ-C30 questions on physical functioning (PF), role functioning (RF), and global health (GH) and quality of life (QOL)

Patient characteristics				T1 (n = 23)			T2 (n = 19)			T3 (n = 23)			T4 (n = 18)		
No.	M/F	Age	LD/ED	GH/QOL	PF	RF	GH/QOL	PF	RF	GH/QOL	PF	RF	GH/QOL	PF	RF
P 12	F	47	LD	83	100	100	100	100	67	66	100	67	83	93	33
P 04	F	50	LD	83	94	67	–	–	–	33	67	83	67	87	83
P 24	F	56	LD	83	100	50	75	93	100	92	93	67	100	100	100
P 17	F	64	ED	83	80	83	83	80	67	83	40	67	#	#	#
P 03	F	64	ED	83	67	33	–	–	–	58	40	0	50	60	33
P 34	F	51	LD	75	100	83	83	100	100	92	100	100	100	100	100
P 15	F	69	LD X*	75	93	83	50	60	67	42	80	67	83	93	100
P 01	F	42	LD X*	66	73	50	–	–	–	50	60	34	83	87	67
P 20	F	44	LD X*	66	87	100	42	73	0	42	80	50	33	47	0
P 32	F	60	LD	66	87	83	100	67	100	83	67	100	83	53	100
P 26	F	59	LD X*	58	42	56	92	93	100	75	92	83	–	–	–
P 08	F	69	ED	50	80	83	33	60	100	75	73	100	83	67	50
P 22	M	55	LD X*	75	73	67	83	75	67	50	60	100	–	–	–
P 21	M	69	ED	75	47	50	83	60	83	83	87	100	83	93	83
P 10	M	46	LD X*	66	87	67	83	87	67	83	93	83	83	93	83
P 09	M	66	ED	66	100	50	83	92	0	92	67	67	75	73	100
P 16	M	68	LD X*	66	67	100	75	73	67	58	67	67	75	87	100
P 18	M	72	ED	66	73	67	50	67	17	50	60	100	66	67	100
P 29	M	63	ED	50	80	50	50	80	67	58	83	33	75	87	100
P 27	M	69	LD	50	53	33	66	53	50	66	67	50	66	80	83
P 13	M	72	ED	50	53	33	66	60	67	66	60	67	42	67	33
P 02	M	57	ED	42	40	33	–	–	–	58	67	100	#	#	#
P 14	M	39	LD	33	53	0	66	73	83	58	60	67	#	#	#
All patients				GH/QOL	PF	RF	GH/QOL	PF	RF	GH/QOL	PF	RF	GH/QOL	PF	RF
Mean score				66	75	62	72	76	67	66	72	72	68	80	75

Small-cell lung cancer patients (n = 23), limited (LD) and extensive (ED) disease receiving 1st line chemotherapy were interviewed at equivalent points in treatment: at start of chemotherapy (T1), 4 weeks later (T2), at end of chemotherapy (T3), and 6 weeks later (T4). Three patients died before T4 (#). Scores range from 0 to 100; higher scores represent a higher level of functioning. Seven patients with limited disease were treated with chemotherapy and radiation therapy (LD X*)

'a little' and explained that they had already been limited for a longer time.

[These examples suggest that limitations were considered as 'normal' (MW)].

Item 2: Do you have any trouble taking a long walk? For many patients this question was difficult to answer, because they had not taken long walks during the previous weeks. Of 23 patients, 11 first struggled with the definition of 'long' (which differed from 10 min to 2 h or from 500 m to 10 km) and then they tried to guess their limitations: *"What is a long walk, 5 km…? I never go for a long walk, I don't like serious walking, but I think I would be limited…a little?"* [M, age 68, T1]. The other 12 mentioned a recent walking experience, and most of them tried to guess: *"I haven't had a walk in the past weeks, so I don't know, but I walked with my son through the corridors in the hospital and that went fine. So, I haven't tried…a little?"* [F, age 69, T1], and two used the same strategy as mentioned earlier (i.e., I didn't walk, so I don't have any trouble).

Item 3: Do you have any trouble taking a short walk outside of the house? Most patients seemed to have an image of a short walk in their mind and circled without hesitation a response category. Others remembered a recent experience: *"A little, I've been to the shops at the end of the street and that was enough for me"* [F, age 64, T1].
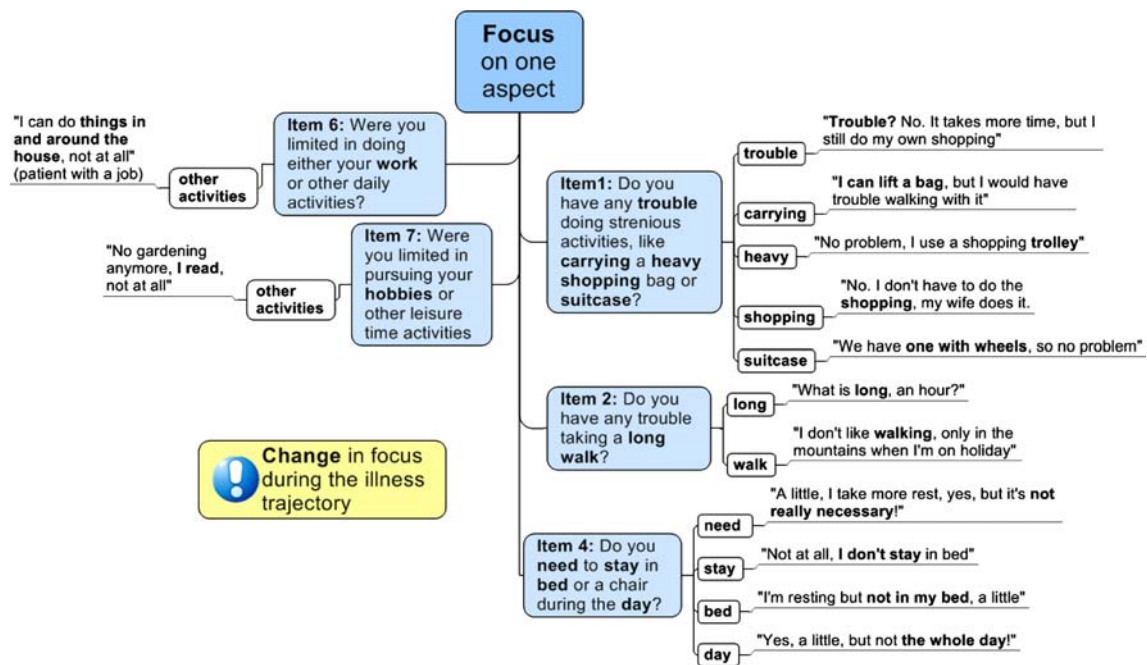
Fig. 1 Examples of response strategies used to answer question 1, 2, 4, 6, and 7 of the EORTC-QLQC30 questionnaire. These strategies and change in the use of a certain strategy over time may explain why patients do not report the deterioration in physical and role functioning that would objectively be expected

Item 4: Do you need to stay in bed or a chair during the day? Many patients answered this question by laying the stress on one word in particular. This resulted in higher levels of functioning than would be expected. For example, 7 out of 23 patients took the word 'need' very seriously and did not report any limitation because *"it's not really necessary to stay in bed or on a chair"*. Another patient interpreted the word 'need' as being prescribed by the doctor. Because he did not have such a prescription, his answer was 'not at all', even though he spent most of the day in his bed [M, age 71, T1]. Another six patients interpreted this question with an emphasis on 'staying in bed', which they did not do during the day. Although these patients said that they took a nap or a rest on a regular basis, they did not take sitting on a chair into account. The other half of the patients did not comment on this item, but just circled the response category of their choice.

Item 5: Do you need help with eating, dressing, washing yourself, or using the toilet? One older male patient considered the help that he received washing and dressing as normal under the circumstances, and said: "No not at all, I can do it by myself if I want to, I don't really need help, but at the moment I have less energy, and help from my wife makes it easier for me" [M, age 72, T1]. Only two patients needed help with eating, dressing, washing, and using the toilet, all the other patients circled the response category of their choice without further comments, or saying that it was self-evident that they were not limited.

Item 6: Were you limited in doing either your work or other daily activities? Most of the retired male patients (*n* = 6) answered with 'a little' or 'not at all', arguing that they were not working anymore and therefore not limited. All female patients who did not have a job (*n* = 8) reported during the first interview the limitations they experienced in house-work. The other nine who worked had taken sick leave, and only three of them (i.e., two male patients with their own business and one female patient with a part-time job) were actually thinking about their job and reported limitations. The other six ignored their work and focused on activities in and around the house: *"I can do some work in the house, like sweeping the floor, so I'm not limited"* [M, age 46, T1].

[These examples suggest that not working was considered as normal under the circumstances (MW)].

Item 7: Were you limited in pursuing your hobbies or other leisure time activities? During the first interview this question was answered with 'very much' in some cases when pursuing hobbies was not possible: "I always played billiards with my friends. I could play at the moment but I don't, because then I'll drink a couple of beers and that's not a good idea now" [M, age 39, T1]. Other patients reasoned as mentioned earlier (i.e., I didn't do it, so I'm not limited): "I've not played the piano in the past few weeks…no, not at all" [M, age 68, T1]. [The answer suggests that the patient is feeling physically able to play, but that he does not take into account the fact that his illness might be the reason for not playing (MW)].

Item 26: How would you rate your overall health during the past week? Item 27: How would you rate your overall quality of life during the past week?

Most patients circled a number between one and seven in response to these items without much thinking aloud. After they had completed these two questions the interviewer asked them how they interpreted the questions. Most patients defined 'overall health' as how they were actually feeling: *"Overall health…, you're not a healthy person of course. Well it's simple, you're terminally ill, but you're not really feeling sick so…a 5"* [M, age 39 at T1]. In evaluating 'overall quality of life', a few patients had just circled a number that they thought appropriate, without knowing what QoL meant. Most patients defined 'overall quality of life' as being able to do the things they want to do: *"I can't do things as usual, watching television, reading my paper, and going to the bookshop. Walking at this moment is not possible…a 4"* [M, age 72 at T1].

Change in patients' answers

The variation in the interpretation and evaluation of the different items described above was not only found between patients, but also in the individual patient over time. Change occurred especially in the sampling of experiences when evaluating problems and limitations concerning items 2, 6, and 7. These changes sometimes resulted in scores suggesting that a patient was functioning better than actually was the case, but we saw the opposite as well. The following examples illustrate these variations in patients' answers.

Item 2: Do you have any trouble taking a long walk? Nina was a women aged 69 (all names are pseudonyms). Her scores indicated an equal level of functioning at both assessments, which contrasted with her described ability to walk at 6 weeks after her chemotherapy (T4) compared to 1 month after the start of her treatment (T2).

[T2] That's very difficult, to the shopping mall, 450 m, quite a bit.
[T4] A long walk, 2 km, I walk too fast, it's my own fault, quite a bit.

[Her answers suggest that she changed her interpretation of a long walk by adjusting the distance. This change can also be interpreted as evidence of response shift, i.e., a change in standards (MW)].

Item 6: Were you limited in doing either your work or other daily activities? Ralph, married with children (age 46), answered this question on both occasions with 'a little'. His scores suggest no change in role functioning. But his concurrent think-aloud texts suggest an improvement.

[T2] A little, it depends how I'm feeling. If I have a good day, I can take on the whole world. Vacuum cleaning, my motorbike, my car".
[T4] A little, the first day back at work again, the tension having to tell everyone the same story over and over again, but of course I feel much better than I did 6 weeks ago.

[At T2 he ignored his job suggesting that taking sick leave was normal. Six weeks after the course he was actually thinking about his first day at work and described a much better state of health as before. Recalling different experiences over time suggest that there was a change in the patient's perspective on "what was considered as normal under the circumstances"(MW)].

Item 7: Were you limited in pursuing your hobbies or other leisure time activities? Like most patients, John (age 69) felt 'quite a bit' limited in pursuing his hobby (gardening) at the first interview (T1). In his next interview, his answer (i.e., 'a little') to the same question suggested an improvement, while his concurrent think aloud shows no improvement at all.

[T1] "My hobby is working in the garden, that's very difficult, quite a bit".
[T2] "I'm reading at the moment. Gardening is not possible anymore, a little".

[At both occasions he was thinking of his hobby (i.e., gardening). However, his answer at the second interview suggests an adjustment to his new situation (MW)].

Item 26: How would you rate your overall health during the past week? Item 27: How would you rate your overall quality of life during the past week?

We found no evidence of change in the patients' definition of overall health. Overall health was consistently interpreted as "how I'm feeling" (e.g., *"Yes, you can't miss it. My health is a 7, it feels like a 7"* [F, age 57, T2]. Neither did we find any change in the interpretation of overall quality of life (i.e., "being able to do the things I want to do"), even in the case of a reoccurence of the cancer and brain metastases (e.g., *"it depends on my contacts…that I can do my own things…a 5 is to low…it has to be a 6"* [F, age 47, T4].

Discussion

The mean values of GH/QoL, physical (PF) and role (RF) functioning scales during the treatment trajectory suggest that SCLC patients had a good overall QoL and were functioning largely without limitations, but we know from interviews that these scores do not accurately reflect QoL

in SCLC patients. We found that the patients evaluated QoL items concerning functioning against a standard of what they perceived as normal under the circumstances and that this standard changed when the circumstances and their health changed. Our results also suggest that patients redefined what is important and what is no longer important (e.g., work or hobbies that used to be important before the diagnosis) through the course of treatment. These findings are consistent with the occurrence of response shifts during the illness trajectory.

Observing the QoL self-assessment by means of the think-aloud procedure has provided us with insight into the black box of what actually happens in repeated QoL measurement. We have lifted the lid just for a little by listening to the patient who is filling in the questionnaire. The physical and role functioning items appeared to have several different interpretations, enabling a patient to give the impression of performing well under the circumstances. Just by ignoring certain activities that are mentioned in a question, or by taking a question literally, patients are able to maintain reasonable levels of functioning, if we are to believe the EORTC data. Through this behaviour it seems that, at a subconscious level, the patients are distancing themselves from the meaning behind the question, i.e., measuring the impact of treatment and disease on their functioning. If this is the case, the patients are presenting their situation more positively than it actually is. Previously [26] we found that, in addition to the response shift phenomena, self-presentation is also a coping mechanism that can explain discrepancies in the measurement of fatigue. The aim of this study was not to investigate whether patients present a more positive image of themselves than they experience in their day-to-day life, but was predominantly to observe how patients interpret and answer the questions, and whether they change their way of answering. We found that the patients did not deal with the questions in the way that the researchers had intended. At face value, some questions were also found to be unimportant, not applicable to the patients' situation, or not relevant at a certain point in time. These findings are in line with Mallinson's suggestion that problems may arise when response options do not quite fit the questions, because there are no such response options as 'I don't do this' or 'I don't know' [37]. On the other hand, the answers of patients who used the strategy 'I can't do this anymore, so I'm not limited' can also be interpreted as evidence of response shift, i.e., reprioritization. Patients make good and legitimate use of the opportunity the QoL instrument provides to adjust the question to their own situation and consequently present an image of not being as limited in functioning as one would expect.

Our findings confirm what has been stated by Rapkin and Schwartz [38], who acknowledge that counter-intuitive findings might be explained by change in the appraisal process: "QoL assessment induces a frame of reference that depends upon the meanings an individual attaches to questions. In responding to items, individuals necessarily sample specific experiences within their frame of reference, and each sampled experience is judged against relevant, subjective standards of comparison. To arrive at a QoL score, individuals must apply some combinatory algorithm to summarize their evaluation of relevant experiences and formulate a response". Changes in the perspective of SCLC patients during their treatment trajectory, changes in their sampling of experiences when evaluating an item, and in the use of standards of comparison explain the variance in QoL measurement at the individual level. Rapkin and Schwartz propose a psychometric model of appraisal that questions the existing methods for establishing the reliability and validity of QoL assessment tools, and they recommend that the assessment of appraisal should be integrated into QoL research and clinical practice. The results of our study show that knowledge about how patients change their interpretation of questions is useful in interpreting QoL data. Therefore, we agree with Rapkin and Schwartz that research is needed to learn more about the appraisal process. Our findings show how deficient items on this questionnaire are in terms of being able to capture the complex response pattern that respondents engage in when filling out the questionnaire. If the items on role and physical functioning and the response options provided lack the sensitivity or specificity to capture the actual responses of the respondents, then this offers an explanation of what has been discussed as discrepant results. Following this line of argument positing a response shift and an appraisal process will have to wait to see if they are present after questions are (re)constructed so as to properly capture the response patterns of the respondents. More care should be taken in assessing the quality of a set of questions to identify problems that result from mismatch between the 'theory' underlying the questions and features of a respondents actual behaviour and biography (29).

Conclusion

Terminally ill patients evaluate their QoL on questionnaires differently than they describe their QoL following what we know to be debilitating chemotherapy treatment. Changes in the appraisal process due to differences in the frame of reference, the sampling of specific experiences, and the standards of comparison used when rating items explain how reported levels of physical and role functioning and QoL are sustained under deteriorating physical conditions. Background knowledge about the illness trajectory and appraisal processes is therefore relevant for the interpretation of QoL outcomes.

References

1. Llewellyn-Thomas, H. A., Thiel, E. C., & Clark, R. M. (1989). Patients versus surrogates: Whose opinion counts on the ethics review panels? *Clinical Research, 37*, 51–55.
2. O'Connor, A. M. (1989). Effects of framing and level of probability on patients' preferences for cancer chemotherapy. *Journal of Clinical Epidemiology, 42*(2), 119–126.
3. Bremnes, R. M., Andersen, K., & Wist, E. A. (1995). Cancer patients, doctors and nurses vary in their willingness to undertake cancer chemotherapy. *European Journal of Cancer, 31A*, 1955–1959.
4. Slevin, M. L., Stubbs, L., Plant, H. J., Wilson, P., Gregory, W. M., & Armes, P. J., et al. (1990). Attitudes to chemotherapy: Comparing views of patients with cancer with those of doctors, nurses, and general public. *BMJ, 300*, 1458–1460.
5. Gralla, R. J. (2004). Quality-of-life considerations in patients with advanced lung cancer: Effect of topotecan on symptom palliation and quality of life. *Oncologist, 9*(Suppl. 6), 14–24.
6. Groen, H. J. M., The, Sanderman, R., De Vries, E. G. E., & Mulder, N. H. (1995). Quality of life in lung cancer patients treated by radiation with and without carboplatin. Thesis Groen HJM. *Exploration of new therapy for lung cancer.*
7. Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., & Duez, N. J., 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. *Journal of the National Cancer Institute, 85*(5), 365–376.
8. Wisloff, F., Eika, S., Hippe, E., Hjorth, M., Holmberg, E., & Kaasa, S., et al. (1996). Measurement of health-related quality of life in multiple myeloma. Nordic Myeloma Study Group. *British Journal of Haematology, 92*(3), 604–613.
9. Andrykowski, M. A., Brady, M. J., & Hunt, J. W. (1993). Positive psychological adjustment in potential bone marrow transplant recipients: Cancer as a psychosocial transition. *Psycho-Oncology, 2*, 261–276.
10. Bach, J. R., & Tilton, M. C. (1994). Life satisfaction and well-being measures in ventilator assisted individuals with traumatic tetraplegia. *Archives of Physical Medicine and Rehabilitation, 75*(6), 626–632.
11. Daltroy, L. H., Larson, M. G., Eaton, H. M., Phillips, C. B., & Liang, M. H. (1999). Discrepancies between self-reported and observed physical function in the elderly: The influence of response shift and other factors. *Social Science & Medicine, 48*(11), 1549–1561.
12. Kagawa-Singer, M. (1993). Redefining health: Living with cancer. *Social Science & Medicine, 37*(3), 295–304.
13. Padilla, G. V., Mishel, M. H., & Grant, M. M. (1992). Uncertainty, appraisal and quality of life. *Quality of Life Research, 1*(3), 155–165.
14. Groenvold, M., Fayers, P. M., Sprangers, M. A., Bjorner, J. B., Klee, M. C., & Aaronson, N. K., et al. (1999). Anxiety and depression in breast cancer patients at low risk of recurrence compared with the general population: A valid comparison? *Journal of Clinical Epidemiology, 52*(6), 523–530.
15. Sneeuw, K. C., Aaronson, N. K., Sprangers, M. A., Detmar, S. B., Wever, L. D., & Schornagel, J. H. (1997). Value of caregiver ratings in evaluating the quality of life of patients with cancer. *Journal of Clinical Oncology, 15*(3), 1206–1217.
16. Sprangers, M. A. G., & Aaronson, N. K. (1992). The role of health care providers and significant others in evaluating the quality of life of parents with chronic disease: A review. *Journal of Clinical Epidemiology, 45*(7), 743–760.
17. Breetvelt, I. S., & Van Dam, F. S. (1991). Underreporting by cancer patients: The case of response-shift. *Social Science & Medicine, 32*(9), 981–987.
18. Frick, E., Borasio, G. D., Zehentner, H., Fischer, N., & Bumeder, I. (2004). Individual quality of life of patients undergoing autologous peripheral blood stem cell transplantation. *Psycho-Oncology, 13*(2), 116–124.
19. Herschbach, P. (2002). Das zufriedenheitsparadox in der Lebensqualitätsforschung. Wovon hängt unser Wohlbefinden ab? *Psychotherapie, Psychosomatik, Medizinische Psychologie, 52*, 141–150.
20. Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine, 48*(11), 1507–1515.
21. Wilson, I. B. (1999). Clinical understanding and clinical implications of response shift. *Social Science & Medicine, 48*(11), 1577–1588.
22. Carver, C. S., & Scheier, M. F. (2000). Scaling back goals and recalibration of the affect system are processes in normal adaptive self-regulation: Understanding 'response shift' phenomena. *Social Science & Medicine, 50*(12), 1715–1722.
23. Sprangers, M. A. (1996). Response-shift bias: A challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treatment Reviews, 22*(Suppl. A), 55–62.
24. Westerman, M. J., Hak, T., The BAM, Groen, H., & van der Wal, G. (2006). Problems eliciting cues in SEIQoL-DW: Quality of life areas in small-cell lung cancer patients. *Quality of Life Research, 15*(3), 441–449.
25. Westerman, M. J., Hak, T., The BAM, Echteld, M. A., Groen, H. J., & van der Wal, G. (2007). Change in what matters to palliative patients: Eliciting information about adaptation with SEIQoL-DW. *Palliative Medicine, 21*(7), 581–586.
26. Westerman, M. J., The BAM, Sprangers, M. A., Groen, H. J., Van der Wal, G., & Hak, T. (2007). Small-cell lung cancer patients are just 'a little bit' tired: Response shift and self-presentation in the measurement of fatigue. *Quality of Life Research, 16*(5), 853–861.
27. Cox, K. (2003). Assessing the quality of life of patients in phase I and II anti-cancer drug trials: Interviews versus questionnaires. *Social Science & Medicine, 56*(5), 921–934.
28. Bergman, B., Aaronson, N. K., Ahmedzai, S., Kaasa, S., & Sullivan, M. (1994). The EORTC QLQ-LC13: A modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC Study Group on Quality of Life. *European Journal of Cancer, 30A*(5), 635–642.
29. Hak, T., Van der Veer, K., & Jansen, H. (2004). The Three-Step Test-Interview (TSTI): An observational instrument for pre-testing self-completion questionnaires. 2004. Rotterdam: Erasmus Research Institute of Management. ERIM Report ERS–2004–029–ORG.
30. O'Boyle, C. A., Hickey, A., Joyce, C. R. B., Browne, J. P., O'Malley, K., & Hiltbrunner, B. (1993). *The schedule for the evaluation of individual quality of life (SEIQOL): Administration manual.* Dublin: Royal College of Surgeons in Ireland, Department of Psychology.

31. O'Boyle, C. A., McGee, H., & Joyce, C. R. B. (1994). Quality of life: Assessing the individual. *Advances in Medical Sociology, 5*, 159–180.

32. Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine, 48*(11), 1531–1548.

33. Sprangers, M. A. (1988). *Response shift and the retrospective pretest*. The Hague: SVO.

34. Sprangers, M. A., Van Dam, F. S., Broersen, J., Lodder, L., Wever, L., & Visser, M. R., et al. (1999). Revealing response shift in longitudinal research on fatigue–the use of the thentest approach. *Acta Oncologica, 38*(6), 709–718.

35. Mays, N, & Pope, C. (2000). Qualitative research in health care. Assessing quality in qualitative research. *BMJ, 320*, 50–52.

36. Yin, R. K. (1989). *Case study research: Design and methods*. London: Sage.

37. Mallinson, S. (1998). The Short-form 36 and older people: Some problems encountered when using postal administration. *Journal of Epidemiology and Community Health, 52*(5), 324–328.

38. Rapkin, B. D., & Schwartz, C. E. (2004). Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health and Quality of Life Outcomes, 2,*14.