RESPONSIVENESS OF THE EQ-5D IN BREAST CANCER PATIENTS IN THEIR FIRST YEAR AFTER TREATMENT

Merel L Kimman*1,2, Carmen D Dirksen3, Philippe Lambin1,2 and Liesbeth J Boersma1,2

Address: 1MAASTRO Clinic, School for Oncology and Developmental Biology (GROW), Maastricht University Medical Centre, Maastricht, the Netherlands, 2Department of Radiation Oncology (MAASTRO), Maastricht University Medical Centre, Maastricht, the Netherlands and 3Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Centre, Maastricht, the Netherlands

Email: Merel L Kimman* - merel.kimman@maastro.nl; Carmen D Dirksen - c.dirksen@mumc.nl; Philippe Lambin - philippe.lambin@maastro.nl; Liesbeth J Boersma - liesbeth.boersma@maastro.nl

* Corresponding author

Abstract

Background/aim: The EQ-5D is a generic health-related quality of life (HRQoL) measure that is used for the purpose of economic evaluations of health interventions. Therefore, it has to be responsive to meaningful changes in health in the patient population under investigation. The aim of this study was to investigate the responsiveness of the EQ-5D in breast cancer patients in their first year after treatment.

Methods: The subscale global health of the disease-specific HRQoL measure EORTC QLQ-C30 was used as a reference instrument to determine meaningful changes in health and identify subgroups of patients: patients reporting a moderate-large deterioration, small deterioration, a small improvement, moderate-large improvement, or no change in health status. Responsiveness was evaluated by calculating standardized response means (SRMs) in the five subgroups of patients and performing analysis of variance procedures. The two HRQoL measures were filled out two weeks and one year after finalizing curative treatment for breast cancer (n = 192).

Results: The EQ-5D was able to capture both improvements and deteriorations in HRQoL. SRMs of the EQ VAS and EQ-5D Index were close to zero in the subgroup reporting no change and increased and decreased adequately in the subgroups reporting small and moderate changes. Additional analysis of variance procedures showed that the EQ-5D was able to differentiate between subgroups of patients with no change and moderate-large deterioration or improvement in health.

Conclusion: The EQ-5D seems an appropriate measure for the purpose of economic evaluations of health intervention in breast cancer patients after treatment.

Trial registration: Current Controlled Trials ISRCTN74071417.
Introduction
With an estimated 1.15 million new cases worldwide each year and a relatively good prognosis, breast cancer is the most prevalent cancer in the world today [1]. After curative treatment for breast cancer, women attend frequent follow-up visits to be examined for possible local or regional recurrence or a second primary breast tumor, and to receive psychosocial support [2,3]. However, no strong evidence exists that regular follow-up is effective with regard to disease free survival or overall survival [4-6], or in providing psychosocial support [7,8]. Hence, the assessment of outcomes like patient satisfaction and health-related quality of life (HRQoL) is common practice in clinical oncology trials investigating alternative follow-up strategies and psychosocial interventions for breast cancer survivors [9-18]. Given the high prevalence of breast cancer and budget constraints in health care, it is also important to understand the impact of alternative strategies on economic outcomes. Therefore, clinical trials are increasingly incorporating generic HRQoL measures, such as the EQ-5D, for the purpose of economic evaluations [19]. The EQ-5D is a standardized multi-dimensional health state classification system. It generates a single index score for each health state [20]. Index scores, in turn, can be used to calculate quality adjusted life years (QALYs), which is the most preferred summary outcome measure in economic evaluations [21].

A substantial and growing body of literature regarding the usefulness of the EQ-5D in cancer has emerged, supporting its validity and reliability [19]. However, the responsiveness of the EQ-5D, defined as its ability to capture true underlying changes in the patients' health status over time [22], is highly dependent on patient population and setting. In comparison with disease-specific instruments, the responsiveness of the EQ-5D was found to be comparable in one study [23], but more often it is found to be less responsive than disease-specific instruments [24-27]. Hence, the usefulness of the EQ-5D may be limited if it is not able to detect changes in health status in the patient population under investigation.

To our knowledge, the responsiveness of the EQ-5D has not yet been examined in breast cancer patients after treatment. Therefore, we use data from a randomized clinical trial investigating several follow-up strategies for curatively treated breast cancer patients [10] to address whether the EQ-5D is responsive to changes in HRQoL in a population of breast cancer patients in their first year after treatment.

Methodology
Study population
Participants were enrolled in a randomized clinical trial investigating the cost-effectiveness of nurse-led telephone follow-up and a short educational group program after curative treatment for breast cancer (MaCare trial, ISRCTN 74071417) [10]. Patients in the trial were all female, treated for breast cancer with curative intent, and had no concomitant tumors or comorbidity requiring hospital visits. There were no age restrictions. Patients were included in the trial after finalizing treatment and after giving written informed consent. Treatment included surgery and/or radiotherapy and/or chemotherapy. Follow-up appointments took place at three, six, nine and twelve months after treatment. For the purpose of studying the responsiveness of the EQ-5D, patients who had had their twelve months follow-up were eligible. The EQ-5D and the disease-specific EORTC QLQ-C30 were sent to patients at home two weeks after the end of treatment (T0) and twelve months after treatment (T1). Of 220 eligible patients, 29 patients failed to complete both instruments at both measurements due to either random missings within the instruments (n = 19) or because they were a study drop-out (n = 10). A total of 192 patients were therefore included in the analysis. Their demographic and clinical characteristics can be found in table 1. Patients were analyzed regardless of follow-up strategy in the trial.

The MaCare trial was approved by the Independent Ethics Committee of MAASTRO Clinic.

HRQoL Instruments
EQ-5D
The EQ-5D is a short generic health-related quality of life instrument that consists of two parts: a self-classifier and a Visual Analogue Scale (EQ VAS). The self-classifier comprises five items relating to problems in the following domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression [20]. Each domain has three levels, namely, "no problems", "some problems" and "severe problems". Combinations of these categories define a total of 243 health states. Dolan et al [28] have presented 42 of these health states to approximately 3000 members of a representative sample of the UK general population, which were valued using the time-trade-off (TTO) technique. Based on these valuations, for each health state a utility score can be deducted, called the EQ-5D Index score. These EQ-5D Index scores may vary between -0.59 (worst health) and 1.00 (perfect health). On the EQ VAS respondents can indicate their overall self-perceived health state on a scale ranging from 0 to 100, where 0 is equivalent to the worst imaginable health state and 100 is equivalent to the best imaginable health state.

EORTC QLQ-C30
The EORTC QLQ-C30 from the European Organization for Research and Treatment of Cancer is a self-administered disease-specific HRQoL questionnaire and is vali-
dated for oncology clinical research [29-31]. It has also been validated [32] and found to be responsive [33] specifically in breast cancer patients and is widely used in breast cancer research investigating HRQoL after treatment [34-40]. The HRQoL questionnaire consists of 30 items. After transformation, the EORTC QLQ-C30 has several multi-item functional subscales (e.g. physical, emotional functioning), multi-item symptom scales (e.g. fatigue, pain), a global health subscale, and single items to assess symptoms (e.g. sleep disturbance). Scores on the functional and global health scales range from 0 to 100, where a higher scale score represents a higher level of functioning and therefore HRQoL.

**Analyses of responsiveness**

To assess the responsiveness of the EQ-5D three steps were taken, following recommendations recently published by Revicki et al (2008). First, a criterion, or anchor, that is related to the measure under investigation, was selected to identify whether patients had changed (either improved or worsened) over time. Second, when the relationship between the anchor and EQ-5D was confirmed, patients were classified into subgroups according to changes in their health status. Third, to examine responsiveness, statistical indicators for change were calculated and analysis of variance procedures were performed.

**Step 1: Selecting an anchor; global health of the EORTC QLQ-C30**

Selecting anchors should be based on criteria of relevance for the disease indication, clinical acceptance and validity, and evidence that the anchors have some relationship with the measure under investigation [41]. For this study, the subscale global health of the EORTC QLQ-C30 was proposed as a criterion for clinical change. The global health subscale consists of two items: (1) How would you rate your physical condition during the past week? and; (2) How would you rate your overall quality of life during the past week?

Correlations between global health scores and the EQ-5D Index and EQ VAS were calculated to examine whether the anchor was acceptable [41]. It is recommended that 0.30–0.35 is used as a correlation threshold to define acceptable association between an anchor and a change score on the HRQoL outcome measure [41].

**Step 2: Classifying patients into subgroups**

Change scores on global health of the EORTC QLQ-C30 were used to identify subgroups of patients. In an analysis of the clinical significance of changes in HRQoL, Osoba et al (1994) showed that patients judge a change between 5–10 on the global health scale of the EORTC QLQ-C30 to be small, between 10–20 to be moderate, and more than 20 to be large [33,42]. Consequently, a change smaller than 5 points was considered to be no change. Taking into account both deteriorations and improvements, this results in a maximum of 7 subgroups.

**Step 3: Examining responsiveness**

Responsiveness to change was evaluated using a statistical indicator, the standardized response mean (SRM). The SRM is the change in score divided by the standard deviation of the change in score. It is independent of sample size and widely used today [43]. SRMs were calculated for the EQ-5D Index and EQ VAS, for all subgroups of patients. Scores were interpreted using benchmarks for effect sizes: 0.20 through 0.49 was interpreted as small, 0.50 through 0.79 as moderate and $\geq 0.80$ as large [44].
Additionally, analysis of variance, with Games Howell post hoc procedures, was performed to compare the mean change scores on the EQ-5D Index and EQ VAS between the 'no change' subgroup and the other subgroups identified in step 2.

**Results**

**Step 1. Selecting an anchor**
The global health scale of the EORTC QLQ-C30 correlated to the change scores of the EQ-5D Index and EQ VAS \( (r = 0.423 \text{ and } r = 0.634 \text{ respectively}) \). Hence, global health was found to be an appropriate anchor and was used to classify subgroups.

**Step 2. Classifying patients into subgroups**
After twelve months, 6 patients (3%) reported a large deterioration on global health, 17 (9%) reported a moderate deterioration, 14 (7%) reported a small deterioration, 55 (28%) reported no change, 28 (16%) reported a small improvement, 32 (17%) a moderate improvement and 40 (21%) reported a large improvement on global health.

Due to a relatively small number of patients reporting a moderate or large deterioration, it was decided to create one subgroup for patients with both moderate and large deteriorations ('moderate-large deterioration') and, for easy comparison, also one subgroup for both moderate and large improvements ('moderate-large improvement'). Hence, five subgroups were identified, classifying patients reporting a (1) moderate-large deterioration \((n = 23)\), (2) small deterioration \((n = 14)\), (3) no change \((n = 55)\), (4) a small improvement \((n = 28)\) and (5) moderate-large improvement in health status \((n = 72)\).

**Step 3. Examining responsiveness**
Mean baseline scores, scores at the twelve month measurement and change scores are presented for all HRQoL measures in table 2. The EQ VAS and EQ-5D Index both moved in the expected direction, indicating negative changes (deterioration) in the subgroups reporting deterioration on global health of the EORTC QLQ-C30 and positive changes (improvements) in the subgroups reporting improvements on global health. Accordingly, only a minor change on the EQ VAS and no change on the EQ-5D Index were reported in the no change subgroup of the EORTC QLQ-C30.

To examine responsiveness, SRMs were calculated for the EQ-5D Index and EQ VAS (table 2). In the subgroup of patients whose global health had not changed, accordingly, neither the SRM of the EQ-5D Index, nor of the EQ VAS indicated an effect. SRMs of the EQ-5D Index for the subgroups indicating a small deterioration or small improvement were too small \((i.e. \text{SRM} < 0.20)\) to be considered as an effect. In contrast, SRMs of the EQ VAS indicated a small effect in these subgroups. SRMs of the subgroups with moderate and large improvements or deteriorations in global health indicated a moderate effect on the EQ-5D Index \((i.e. \text{SRM} > 0.50)\) and a large effect on the EQ VAS \((i.e. \text{SRM} > 0.80)\).

Analysis of variance procedures were performed to evaluate whether the EQ-5D could discriminate between the five subgroups (table 3). Results indicated that when the EQ-5D Index score was used as the outcome measure, the subgroup reporting no change on global health differed significantly from the subgroup reporting moderate and large improvements. The subgroups reporting small improvements or a small or moderate and large deteriorations.
tion could not be differentiated from the 'no change' subgroup. The EQ VAS on the other hand was able to discriminate between the 'no change' subgroup and the subgroups reporting a moderate and large improvement and moderate and large deterioration.

Discussion

An increasing number of clinical trials is investigating the effectiveness of follow-up strategies and psychosocial interventions for breast cancer patients after treatment, using HRQoL as an important outcome measure [45,46]. Hence, a good responsiveness of the HRQoL measure used seems essential. Our study showed that the EQ-5D was able to detect both improvements and deteriorations in health. However, according to Cohen’s benchmarks for effect sizes [44], the EQ-5D Index was not responsive to small changes in health. The inability of the EQ-5D Index to detect small changes might be explained by its structure. It is generally acknowledged that more response options lead to a higher responsiveness [26]. The domains of the EQ-5D have only three response levels, making it difficult to pick up small changes in health. In addition, in the subgroup of patients reporting no change and the subgroup reporting no change on global health, baseline scores on the EQ-5D Index were relatively high. These high scores were a result of large proportions of respondents already in the top category of domains of the EQ-5D. This ceiling effect is a well known feature of the EQ-5D and left little room for improvement [47]. A straightforward solution would be to attempt to produce a better, more responsive, generic index measure. Recent studies on an EQ-5D with five response levels for each domain showed increased descriptive power and suggest better discriminatory power [48,49]. Hence a less severe ceiling effect and increased benefit in the detection of small health changes are expected [49]. Unfortunately, an official five-level descriptive system is not yet available.

Additional analysis of variance procedures to investigate responsiveness showed that the EQ-5D Index and the EQ VAS both could not differentiate between subgroups reporting no change and small changes in global health. For the EQ-5D Index this was in accordance with the small SRMs in these subgroups. For the EQ VAS however, the non-significant differences were unexpected, as the SRMs indicated moderate effects. This inability of the EQ VAS to discriminate might be explained by the small number of patients in these subgroups (n = 14 and n = 28 respectively). Analysis of variance procedures and especially post hoc procedures are sensitive to population variances and differences in sample size in subgroups. Hence, with a larger sample size, the EQ VAS might have been able to differentiate between subgroups with no change and small changes in health. This argument also holds true for the EQ-5D Index, which could not discriminate between the 'no change' subgroup and the subgroup reporting a moderate-large deterioration in health (n = 23).

A limitation of this study was that the responsiveness was investigated using a single anchor, while ideally multiple anchors should be used to investigate the responsiveness of an instrument [50]. A clinical variable, such as whether or not a recurrence was detected, would be a suitable second anchor to classify subgroups of patients. However, in the clinical trial from which participants were used for these analyses, only few (< 10) recurrences were reported, and unfortunately, these participants were study dropouts. Hence, an appropriate second anchor was not available. Further research into the responsiveness of the EQ-5D in breast cancer patients should aim to include multiple anchors.

In summary, results of this study showed that the EQ-5D was able to capture both improvements and deteriorations in HRQoL of breast cancer patients after treatment, but small changes in health were not recognized as being meaningful. However, in economic evaluations the EQ-5D is primarily used to measure outcome for QALY analysis rather than measuring HRQoL for clinical purposes. Within the framework of economic evaluations, an incremental cost-effectiveness ratio (i.e. additional cost per QALY gained) is more informative than the difference in HRQoL alone. Therefore, a small difference in the EQ-5D Index might still be meaningful when additional costs for

Table 3: Analysis of variance

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>EORTC QLQ-C30</th>
<th>Mean difference (SE)</th>
<th>p-value</th>
<th>EQ-5D Index</th>
<th>Mean difference (SE)</th>
<th>p-value</th>
<th>EQ VAS</th>
<th>Mean difference (SE)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate-large deterioration (n = 23)</td>
<td></td>
<td>-0.14 (0.07)</td>
<td>.228</td>
<td>-13.88 (2.65)*</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small deterioration</td>
<td></td>
<td>0.01 (0.05)</td>
<td>1.000</td>
<td>-5.78 (3.44)</td>
<td>.470</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No change (n = 55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small improvement (n = 28)</td>
<td></td>
<td>0.02 (0.04)</td>
<td>.984</td>
<td>5.37 (1.93)</td>
<td>.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate-large improvement (n = 72)</td>
<td></td>
<td>0.13 (0.04)*</td>
<td>.006</td>
<td>11.38 (1.97)*</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Subgroups are significantly different at a 0.05 significance level
such a change in HRQoL are very low. Hence, the EQ-5D should indeed be able to pick up relevant changes in health and should be able to differentiate between subgroups of patients to some extent, but cut-off points for effect sizes or discriminative ability are less relevant in the context of economic evaluations.

Conclusion
In this study the responsiveness of the EQ-5D was investigated for its use in economic evaluations of health interventions in breast cancer patients after primary treatment. The EQ-5D was able to detect improvements and deteriorations in health and could discriminate between patients with no change in health and patients with moderate-large changes in health. Therefore, the EQ-5D seems an appropriate HRQoL measure for economic evaluations in breast cancer patients after treatment.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
MK was responsible for the data collection and drafted the manuscript. MK works under direct supervision of CD and LB. PL, CD and LB read and corrected draft versions of the manuscript.

Acknowledgements
This research is funded by the Netherlands Organization for Health Research and Development (ZonMw grant no. 945-04-512, ISRCTN 74071417). The authors would like to thank Karin de Bie for her assistance with the data collection.

References


