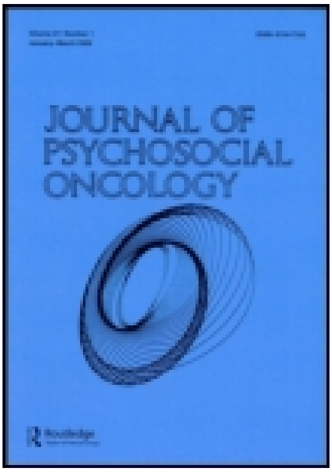

Journal of Psychosocial Oncology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/wjpo20

Summarizing the Fifteen Scales of the EORTC QLQ-C30 Questionnaire by Five Aggregate Scales with Two Underlying Dimensions: A Literature Review and an Empirical Study

Willem A. van der Kloot PhD[a], Kunihiko Kobayashi MD, PhD[b], Kazue Yamaoka PhD[c], Kenichi Inoue MD, PhD[d], Hans W. R. Nortier MD, PhD[e] & Ad A. Kaptein PhD[f]

[a] Institute of Psychology, Leiden University, Leiden, The Netherlands

[b] Department of Respiratory Medicine, Saitama International Medical Center, Hidaka City, Japan

[c] Graduate School of Public Health, Teikiyo University, Tokyo, Japan

[d] Saitama Cancer Center, Saitama, Japan

[e] Department of Clinical Oncology, Leiden University Medical Center, Leiden, The Netherlands

[f] Psychology Unit, Leiden University Medical Centre (LUMC), Leiden, The Netherlands

Accepted author version posted online: 05 May 2014.Published online: 02 Jul 2014.

Journal of Psychosocial Oncology, 32:413–430, 2014

ISSN: 0734-7332 print / 1540-7586 online
DOI: 10.1080/07347332.2014.917139

Summarizing the Fifteen Scales of the EORTC QLQ-C30 Questionnaire by Five Aggregate Scales with Two Underlying Dimensions: A Literature Review and an Empirical Study

WILLEM A. VAN DER KLOOT, PhD
Institute of Psychology, Leiden University, Leiden, The Netherlands

KUNIHIKO KOBAYASHI, MD, PhD
Department of Respiratory Medicine, Saitama International Medical Center, Hidaka City, Japan

KAZUE YAMAOKA, PhD
Graduate School of Public Health, Teikiyo University, Tokyo, Japan

KENICHI INOUE, MD, PhD
Saitama Cancer Center, Saitama, Japan

HANS W. R. NORTIER, MD, PhD
Department of Clinical Oncology, Leiden University Medical Center, Leiden, The Netherlands

AD A. KAPTEIN, PhD
Psychology Unit, Leiden University Medical Centre (LUMC), Leiden, The Netherlands

The intercorrelations among the 15 scales of the 30-item Core version of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire suggest that one may combine (1) the physical functioning and role functioning scales, (2) the emotional functioning and cognitive functioning scales, and (3) the nine symptom scales. Together with the global health/quality of life scale and the social functioning scale, five measures remain. Principal component analysis of those five measures, using data from Japanese and Dutch breast and lung cancer patients, yielded two dimensions: (1) generalized health related quality of life and

Address correspondence to Willem A. van der Kloot, PhD, Institute of Psychology, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands. E-mail: w.a.van.der.kloot@umail.leidenuniv.nl

(2) health-independent psychological well-being. The correlations of these dimensions with the Brief Illness Perception Questionnaire and Karnofsky performance substantiated this interpretation.

KEYWORDS health-related quality of life, EORTC QLQ-C30, illness perception, cancer, categorical principal component analysis

INTRODUCTION

As health-related quality of life (HRQOL) has become a major point of interest in cancer care and research, many instruments have been developed to measure this construct in patients suffering from cancer. One such instrument, the Quality of Life Questionnaire of the European Organization for Research and Treatment of Cancer (the EORTC QLQ-C30), has become one of the standard instruments for measuring HRQOL in patients with any form of cancer (Aaronson, et al., 1993).

In the 30- item EORTC QLQ-C30 (QLQ-C30, for short) respondents obtain scores on 15 scales: global health/quality of life (QL, two items), physical functioning (PF, five items), role functioning (RF, two items), social functioning (SF, two items), cognitive functioning (CF, two items), emotional functioning (EF, four items), fatigue (FA, three items), nausea/vomiting (NV, two items), pain (PA, two items), and the single item scales dyspnoea (DY), insomnia (SL), appetite loss (AP), constipation (CO), diarrhea (DI), and financial difficulties (FI). The scores are made to range from 0 to 100%. Higher scores on QL, PF, RF, SF, CF, and EF indicate better functioning. Higher scores on the nine symptom scales indicate more intense symptoms.

HRQOL is a multidimensional construct pertaining to the physical, mental, and social condition of the patient. The 15 QLQ-C30 scores provide information on many aspects of those conditions. It has been noted that several scales of the QLQ-C30 are interrelated, but that the nature of these relationships is "understudied and not yet clear" (Oerlemans, Mols, Nijziel, Lybeert, & van de Poll-Franse, 2011, p. 1002). It has also been argued that the 15 scales of the QLQ-C30 might be represented by fewer summary measures. In research applications, reducing the number of scores may have the advantages of fewer Type I errors and increased statistical testing power (Gundy et al., 2012; King, Dobson, & Harnett, 1996; McLachlan, Devins, & Goodwin, 1999), greater precision of measurement (Gundy et al., 2012), and—if the measures are aggregated into one score—improved comparability of scores across different instruments (Pagano & Gotay, 2006). Summary scores may also reduce the number of missing data and are more easily used as stratification variables (McLachlan et al., 1999). However, different settings may require different levels of detail. As Gundy et al. (2012) noted "it might sometimes be more useful, particularly in clinical trials, to employ a composite variable measured with greater precision . . . , as opposed to many variables, each measured with less precision" (p. 1608).

Mc Lachlan et al. (1999) stated that summary scores may have advantages in health policy analyses and economic outcome evaluation, in screening, in population health monitoring and subgroup comparison. They noted, however, that

> the level of detail provided by the instrument is important and required for addressing the types of research questions most typically posed in phase III oncology clinical trials. In these studies, researchers are usually interested in the separate subscale scores of a questionnaire. (p. 315)

As King et al. (1996) put it,

> the trade-off between the number of QOL dimensions measured and the statistical power of each one is worth considering for clinical trial applications . . . . It is not relevant, however, in the management of individual patients, where full information takes precedence. (p. 28)

Nevertheless, even in clinical settings, 15 measures may be rather numerous, particularly as the information may be partially redundant. If busy clinicians can focus on a small number of major aspects rather than on many specific details, it becomes easier to monitor a patient, especially over time. A case in point is the frequent use of single-item "thermometers" for—for instance—pain or distress. Having fewer questionnaire scores to consider would not only reduce the clinician's workload, but the use of summary scores based on many QOL items would yield a more reliable measure with greater precision and thus do more justice to the individual patient in a clinical situation.

Therefore, it is a relevant question whether the information of the QLQ-C30 can be represented in a more parsimonious and manageable manner by aggregating subsets of the 15 scores. To answer this question we have conducted a literature review followed by an empirical study.

LITERATURE REVIEW

Many studies have investigated the correlations among subsets of scales and items of the QLQ-C30. In five studies the intercorrelations of the individual items were analyzed to determine the reproducibility of the multi-item scales (Ford, Havstadt, & Kart, 2001; Gotay, Blaine, Haynes, Holup, & Pagano, 2002; Kart & Ford, 2002; McLachlan et al., 1999; Osaba et al., 1994). Five studies explicitly analyzed the intercorrelations and underlying structure of the scales (Boehmer & Luszczynska, 2006; Gotay et al., 2002; Gundy et al., 2012; Ringdal & Ringdal, 1993; van Steen et al., 2002). Seven studies investigated the correlations of subsets of the QLQ-C30 with several scales and items of other instruments (Arraras et al., 2002; Arraras Urdaniz et al., 2008; Henoch, Plone, & Tishelman, 2009; King et al., 1996; Kobayashi et al., 2008; Pagano & Gotay, 2006; Strasser, Müller-Käser, & Dietrich, 2009). A detailed description of those studies is given in the appendix. The main results are summarized in Table 1.

Table 1 shows that many scales are related to each other, though not always in the same combinations. Two frequent findings are the associations between PF and RF and between CF and EF. Therefore it is defendable to combine PF and RF, on the one hand, and CF and EF, on the other hand. Table 1 also shows that there are no stable combinations among the symptom scales. Given the results of Boehmer and Luszczynska (2006), Gundy et al. (2012), and Henoch et al. (2009), it is defendable to combine them into one "symptomatology" indicator. Together with the QL and SF scales, this would lead to five measures instead of the original 15.

EMPIRICAL STUDY

In this study we wanted to examine the dimensionality of the EORTC QLQ-C30 after the 15 original scales had been aggregated—see above—into five scales: QL, SF, the combination of PF and RF, the combination of CF and EF, and the combination of the nine symptom scales. We expected that those five scales would be substantially correlated and could further be aggregated. We hypothesized that the five scales would contain two factors or dimensions: (1) a strong HRQOL factor with high positive loadings of QL, SF, and the PF-RF and CF-EF combinations and a large negative loading of the combined symptoms, and (2) a second factor that discriminates between the PF-RF and the CF-EF combination. We had no a priori expectations about the position of SF. These expectations were tested by means of correlation coefficients and principal component analysis. The interpretation of the resulting components (dimensions, factors) was tested by relating the components to the Brief Illness Perception Questionnaire (B-IPQ; Broadbent, Petrie, Main, & Weinman, 2006) and Karnofsky performance status (Karnofsky, Abelman, Craver, & Burchenal, 1948).

METHOD

Patients

In this international study, the data were obtained from 22 Japanese and 24 Dutch non-small-cell lung cancer patients and 21 Japanese and 22 Dutch patients with breast cancer. Patients completed a questionnaire booklet immediately before their first chemotherapy cycle, one week after their first chemotherapy cycle, and 8 weeks after the start of chemotherapy. The

TABLE 1 Summary of Relationships Among the Scales of the 30 Item Core Version of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) Found in the Literature (Identical Symbols Within a Row Indicate That Scales Belong to the Same Factors, Clusters, or Constructs)

	QL	PF	RF	SF	CF	EF	FA	NV	PA	DY	SL	AP	CO	DI	FI
Construct validity studies															
Ford (2001)	-[a]	●	-	-	-	□	○	-	-	-	-	-	-	-	-
Gotay (2002) Study 2	●	●	●	●	●	●	?	?	?	?	?	?	?	?	?
Kart (2002)	?[b]	●	?	?	◇	○	●	?	?	?	?	?	?	?	?
McLachlan (1999)	○/■	-	○	○	■	■	-	-	-	-	-	-	-	-	-
Osaba (1994)	□	●	●	?	?	?	●/□	?	?	?	?	?	?	?	?
Scale interrelation studies															
Boehmer (2006)	-	●	●	●	●	●	□	□	□	□	□	□	□	□	□
Gotay (2002) Study 3	□	□	□	□	□	□	-	-	-	-	-	-	-	-	-
Gundy et al. (2012)	□	●	●/○	●/○	○	○	●/○	●	●/○	●	●/○	●	●	●	-
Ringdal (1993)	●/□	●	●	□	□	□	●/□	●	●	?	?	●	?	?	?
Van Steen (2002)	●	●	●	●	○	○/□	○	○	●	○	□	○	-	-	-
Studies relating EORTC QLQ-C30 scales to other instruments															
Arraras (2002)	?	●	●	●	?	●	●	?	?	?	?	?	?	?	?
Arraras (2008)	?	●	●	●	?	□	●/□	?	?	?	?	?	?	?	?
Henoch (2009)	-	-	-	-	□	□	○	○	○	◇	□	○	○	○	-
King (1996)	□	●	●	○	?	○	□	?	□	-	-	-	-	-	-
Kobayashi (2008)	?	●	●	●	●	●	-	-	-	-	-	-	-	-	-
Pagano (2006)	●	●	●	●	?	●	●	?	●	?	?	●	?	?	?
Strasser (2009)	-	●	-	-	□	●/□	-	-	-	-	-	-	-	-	-

QL = global health/quality of life; PF = physical functioning; RF = role functioning; SF = social functioning; CF = cognitive functioning; EF = emotional functioning; FA = fatigue; NV = nausea/vomiting; PA = pain; DY = dyspnoea; SL = insomnia; AP = appetite loss; CO = constipation; DI = diarrhea; FI = financial difficulties.
a. - The corresponding scale was not included in the study;
b. ? The corresponding article did not contain information about the scale or information that had no clear interpretation.

international research project was approved by the Medical Ethical Committee of the Leiden University Medical Centre, and by the Internal Review Board of the Saitama International Medical Centre, Hidaka City, Japan. This article is one of a series of publications on this project. Some results have been published in Kaptein et al. (2011) and Kaptein et al. (2013).

Questionnaires

The questionnaire booklets included the QLQ-C30 (Aaronson et al. 1993; Kobayashi et al., 1998) on all three occasions and the B-IPQ (Broadbent et al., 2006) on the first occasion.

The B-IPQ contains eight questions that measure eight dimensions of illness perception: consequences (How much does your illness affect your life?), timeline (How long do you think your illness will continue?), personal control (How much control do feel you have over your illness?), treatment control (How much do you think your treatment can help your illness?), identity (How much do you experience symptoms from your illness?), concern (How concerned are you about your illness?), coherence (How well do you feel you understand your illness?), and emotional response (How much does your illness affect you emotionally? e.g., does it make you angry, scared, upset or depressed?). The responses are measured on a scale of 1 (*not at all*) to 10 (*very much*). For the Dutch and Japanese versions see www.uib.no/ipq.

Physicians rated the Karnofsky performance status (Karnofsky et al., 1948) before the first chemotherapy cycle. The Dutch patients also provided self-ratings of their performance status on all three occasions.

Data Analysis

The QLQ-C30 was scored according to the manual (Fayers et al., 2001). PF and RF were averaged; we label this measure *PRF* for Physical and Role Functioning. EF and CF were also averaged; the resulting variable is labeled *PSY*, for psychological functioning. All symptom scores were averaged as well, yielding a new variable *SYM* (symptomatology).

Pearson product-moment correlations were used to study the relations among the five scales. To control for differences among countries and occasions, pooled within-country-and-occasion correlations were computed. Principal component analysis was used to represent the patients and occasions in a reduced number of dimensions. We used the Categorical Principal Component Analysis program (CATPCA) of SPSS-17 because it can handle missing data and can incorporate variables with different measurement levels (Linting, Meulman, Groenen, & van der Kooij, 2007).

RESULTS

Of the 22 Japanese lung cancer patients 5 (22.7%) were female and 17 (77.3%) were male. The mean age (± Standard Deviation) of these patients was 63.0 ± 6.6 years. The Dutch lung cancer group consisted of 8 (33.3%) female and 16 (66.7%) male patients, with mean age of 63.3 ± 9.7 years. The mean age of the 21 Japanese breast cancer patients was 49.9 ± 9.6 years, and the mean age of the 22 Dutch patients was 46.8 ± 7.8 years.

Two Japanese patients (one lung cancer, one breast cancer) with missing data on all QLQ-C30 questions of Occasions 2 and 3 were omitted from all analyses. Seven Japanese lung cancer patients had completely missing QLQ-C30 data on Occasion 3, which reduced the number of patients on this occasion to 80. Ten of the remaining Japanese and Dutch patients had missing data on one or two QLQ-C30 variables on one of the three occasions. The nonmissing data of the latter patients were included in the analyses. Therefore the data to be analyzed consisted of 254 observations (87 patients on Occasion 1, 87 patients on Occasion 2, 80 patients on Occasion 3).

Intercorrelations

Table 2 contains the pooled within-country-and-occasion product–moment correlation coefficients between the 15 original QLQ-C30 measurements and the five scales explained above, as well as the intercorrelations among the latter scales. The coefficients in this table show that PRF is an adequate representation of PF and RF ($r = .870$ and $r = .955$, respectively). Similarly, PSY summarizes EF and CF quite well ($r = .894$ and $r = .906$, respectively). The symptomatology scale SYM adequately represents FA, AP, PA, SL ($r = .824$, $r = .753$, $r = .748$, $r = .\ 712$), and—to a lesser extent—NV, FI, CO, DY, and DI ($r = .580$, $r = .461$, $r = .454$, $r = .450$, $r = .272$). Table 2 also shows that QL, SF, PRF, PSY, and SYM have substantial correlations with each other. All correlations are in the expected direction: positive coefficients among the functioning scales, and negative coefficients for the correlations between functioning scales and symptoms. These results indicate that the five summary scales of the QLQ-C30 still can be further aggregated.

Principal Component Analysis

The data that were analyzed consisted of 254 observations (87 patients on Occasion 1, 87 patients on Occasion 2, 80 patients on Occasion 3) on five variables (QL, SF, PRF, PSY, SYM). Before running the CATPCA analysis the scores of all variables were discretized such that each category contained approximately 20 observations. The discretized variables were treated as

TABLE 2 Pooled Within-Country-and-Occasion Correlations Between Five Summary Scales and the Original 15 Scales of the 30-Item Core Version of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) and Intercorrelations of the Five Summary Scales

	QL	SF	PRF	PSY	SYM
			Original EORTC-QLC-C30 scales		
QL	1	.577**	.632**	.493**	−.696**
PF	.605**	.476**	.870a**	.462**	−.623**
RF	.571**	.646**	.955a**	.507**	−.638**
SF	.577**	1	.631**	.549**	−.685**
CF	.464**	.533**	.590**	.906a**	−.637**
EF	.422**	.453**	.359**	.894a**	−.516**
FA	−.638**	−.601**	−.756**	−.578**	.824a**
NV	−.339**	−.369**	−.284**	−.372**	.580a**
PA	−.570**	−.517**	−.654**	−.577**	.748a**
DY	−.353**	−.186**	−.413**	−.199*	.450a**
SL	−.470**	−.469**	−.440**	−.530**	.712a**
AP	−.575**	−.448**	−.535**	−.458**	.753a**
CO	−.235**	−.254**	−.246**	−.266**	.454a**
DI	−.139	−.119	−.161	−.121	.272a**
FI	−.323**	−.367**	−.127	−.263**	.461a**
			Summary scales		
QL	1	.577**	.632**	.493**	−.696**
SF	.577**	1	.631**	.549**	−.633**
PRF	.632**	.631**	1	.531**	−.685**
PSY	.493**	.549**	.531**	1	−.642**
SYM	−.696**	−.633**	−.685**	−.642**	1

QL = global health/quality of life; PF = physical functioning; RF = role functioning; SF = social functioning; CF = cognitive functioning; EF = emotional functioning; FA = fatigue; NV = nausea/vomiting; PA = pain; DY = dyspnoea; SL = insomnia; AP = appetite loss; CO = constipation; DI = diarrhea; FI = financial difficulties; PRF = physical and role functioning; PSY = psychological functioning; SYM = symptomatology.
a. Correlations of Summary scales with their constituents.
$^{*}p < 0.05$, $^{**}p < 0.01$ (two-tailed), $df = 145$.

ordinal data, that is, the ranks of the categories and not their exact values provided the core information for the analysis (Linting et al., 2007).

Two principal components were extracted. The first component (eigenvalue: 3.386) explained 67.7% of the variance (VAF). The second component had a substantially smaller eigenvalue (.681 or 13.6%). A separate CATPCA in the Japanese group yielded components with VAFs of 66.1% and 14.8%; in the Dutch group VAFs were 70.4% and 16.0%. These values indicate that a one-component solution is defendable. Nevertheless, we have chosen two components because the second component still accounted for a substantial percentage of the variation and because it was needed to sufficiently represent PSY.

CATPCA yields two sets of important outcomes: (1) component scores (i.e., weighted combinations of the original variables, sometimes called factor scores) for all observations, and (2) loadings for all variables (i.e., correlations

TABLE 3 Loadings of the Five Summary Scales on Two Principal Components

	Component 1	Component 2
QL	.843	−.215
SF	.848	−.042
PRF	.837	−.330
PSY	.679	.724
SYM	−.891	−.003

QL = global health/quality of life; SF = social functioning; PRF = physical and role functioning; PSY = psychological functioning; SYM = symptomatology.

of the original variables with the components; see Table 3). These outcomes are depicted in Figure 1. This so-called biplot (Greenacre, 2010) contains 254 points (3 points for each of 80 patients and 2 points for each of seven patients). The projections of the patients' points on the horizontal and vertical dimensions represent the patients' scores on the components. In addition to the patient points, the biplot contains five arrows that indicate the loadings of the variables. The arrows show that all function scales have high positive loadings on the first (horizontal) component, whereas SYM loads highly in the opposite direction. This first component is interpreted as a generalized

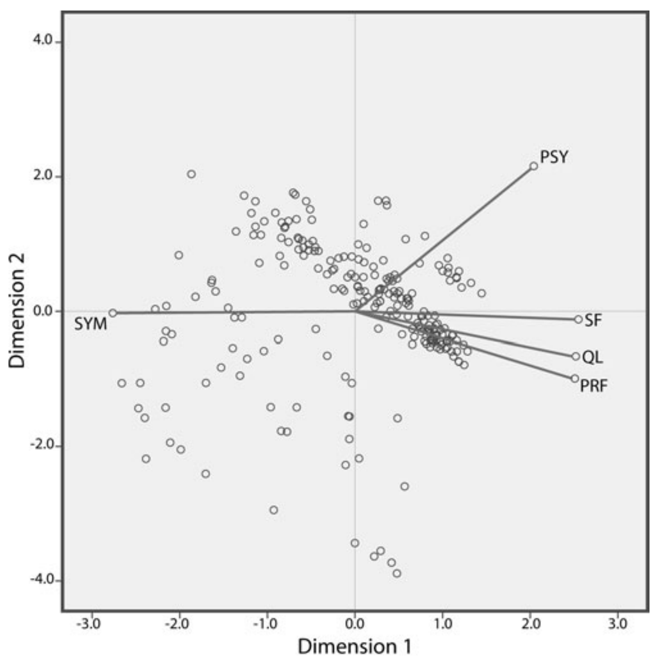

FIGURE 1 Biplot of patients and EORTC QLQ-C30 summary variable in a two-dimensional space. QL = global health/quality of life; SF = social functioning; PRF = physical and role functioning; PSY = psychological functioning; SYM = symtomatology.

HRQOL dimension. On the second (vertical) dimension, only PSY has a substantial loading, although it clearly does not coincide with this dimension. Therefore, this dimension measures those aspects of psychological well-being (or rather, the absence of cognitive and emotional problems) that are not related to the symptomatology and the physical, role, and social components of HRQOL. Therefore, this dimension appears to measure that part of psychological well-being that is not connected to physical health.

Separate CATPCAs of the Japanese and Dutch groups yielded strong first components that we interpreted as generalized HRQOL dimensions. Some differences were found with regard to SF and QL. In the Dutch patients, QL had also a substantial loading on the second dimension, in the opposite direction of PSY. In the Japanese group, the same pattern was found for SF. However, as in both analyses, PRF, QL, and SF are closer to each other than to PSY, we can interpret the principal components as a HRQOL and a dimension of unrelated psychological well-being, both in the Japanese and in the Dutch data.

To validate this interpretation we used the component scores as predictors in regression analyses with gender, age, the Karnofsky ratings by doctors of all patients on Occasion 1, the Karnofsky self-ratings of the Dutch patients on three occasions, and the eight B-IPQ measures as dependent variables. The results, which are displayed in Table 4, show that Karnofsky

TABLE 4 Regression Statistics Describing the Linear Relations Between the Two Principal Components of the Summary Scales of the 30-Item Core Version of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) and Several Supplementary Variables

	Standardized Regression Coefficients		
Variable	Component 1	Component 2	Multiple Correlation Coefficient
Gender (male = 1, female = 2)	.025	−.059	.064
Age	.024	−.041	.046
Karnofsky by doctor	.359**	−.195	.381**
Karnofsky self ratings Occasion 1[a]	.676***	−.037	.673***
Karnofsky self-ratings Occasion 2[a]	.581***	.192	.634***
Karnofsky self-ratings Occasion 3[a]	.746***	−.086	.731***
B-IPQ consequences	−.445***	−.021	.448***
B-IPQ time line	−.249*	.052	.248
B-IPQ personal control	−.168	.079	.176
B-IPQ treatment control	−.080	.030	.083
B-IPQ identity	−.425***	.022	.423***
B-IPQ concern	−.231***	−.184*	.312*
B-IPQ coherence	−.052	.013	.052
B-IPQ emotional response	−.546***	−.212***	.609***

BIPQ = Brief Illness Perception Questionnaire.
[a]Dutch patients only.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

performance status, whether rated by doctors or by the patients themselves, is substantially related to the first component only. The same holds for the B-IPQ variables consequences (influence on one's life), time-line (expected duration of the illness), and identity (physical complaints), which all have significant negative correlations with only the first component. This substantiates the interpretation of the first component as a generalized HRQOL dimension. The B-IPQ variables concern (worry about one's illness) and emotional response (e.g., angry, afraid, upset, depressed) are also significantly (negatively) related to the second component. This supports the interpretation of the second component as a psychological dimension.

DISCUSSION

Review of the literature suggested that several QLQ-C30 scales may be combined: PR and RF into PRF, EF and CF into PSY, and the nine symptom scales into one symptomatology scale SYM. The correlations between PRF, PSY, and SYM and the original 15 QLQ-C30 scales demonstrated that these summary scales are adequate representations of their constituents. The substantial correlations among PRF, PSY, SYM, QL and SF indicated that further reduction was possible.

The configuration found in the principal component analysis confirmed our hypothesis that the QLQ-C30 can be represented in a small number of dimensions: a generalized HRQOL dimension and a psychological dimension. This interpretation was supported by the correlations of the QLQ-C30 components with the dimensions of the B-IPQ and the ratings of Karnofsky performance status.

As mentioned in the Results section, our second or psychological dimension appears to measure a patient's psychological well-being that is independent of her or his physical health. High scores on this dimension could indicate happiness, optimism, confidence, a feeling of coping; low scores could indicate the opposite: unhappiness, pessimism, concern, anxiety, a feeling of having lost one's grip on life, and so on. Scores on this dimension vary independently of physical complaints and general quality of life. In the context of chemotherapy, one could easily imagine a patient whose physical condition and related quality of life is poor, but whose psychological well-being—nevertheless—is positive because being treated brings hope for remission and gratitude for health care providers. Vice versa, patients with few complaints and good general quality of life may have low psychological well-being, for instance because they are devastated by the diagnosis of cancer. Whether our psychological dimension reflects permanent personality characteristics or temporary states is a matter of further research.

Our conclusion that the QLQ-C30 scales can be reduced to a generalized HRQOL and a psychological component is consistent with the physical/mental health model of Gundy et al. (2012), who conducted the most extensive study on the dimensionality of the QLQ-C30. However, in our representation QL coincided with generalized HRQOL, whereas Gundy et al. treated QL as a separate, though correlated, latent factor. We expect that the correlations of QL with the physical and mental health factors of Gundy et al. (coefficients were not reported) are very substantial. It is what they should be if the other scales are really measuring HRQOL.

Our findings might be fruitfully applied, in that attention can be focused on a smaller number of QLQ-C30 measurements (five instead of 15). In some cases one might even represent the QLQ-C30 measures by just two component scores, that is, by two (weighted) combinations of the original scores. For instance, the aggregate score QLCOMP1 = {QL + PF + RF + SF + CF + EF − (FA+ NV + PA + DY + SL + AP + CO + DI + FI)/9} is an excellent approximation of the component scores on Dimension 1 of this study (r = .980). QLCOMP2 = {CE + EF−(QL + RF)/2} is a good approximation of the component scores on Dimension 2 (r = .859). Note: subtracting the mean of QL and RF suppresses the contribution of the physical health dimension from the psychological scales.

We want to emphasize that we do not propose to change or shorten the original EORTC QLQ-30 questionnaire, nor do we recommend that the original scales are scored differently. In fact, we start with the very scales prescribed and scored according to the manual (Fayers et al., 2001). We only suggest that the original scales may be combined afterwards, that is, when employing 15 measurements is impossible or impractical. For instance: in a clinical setting, monitoring five or two instead of 15 measurements may simplify the tasks of clinical practitioners and relieve their workload. Moreover, composite scores of several variables may increase precision. In a research setting, the power of simultaneous statistical tests can be increased by performing fewer tests on fewer variables.

This study has also some bearing on the underlying structure of the B-IPQ. The regression coefficients of Table 4 indicate that time-line, consequences, and identity might to some extent be redundant. The same appears to hold for concern and emotional response. Personal control, treatment control, and coherence may, on the other hand, be independent factors.

The generalizability of the above results has some limitations. First, they are based on the data of a relatively small number of patients. However, one could argue that 87 patients is a fair sample size to study the relationships among five correlated variables, and that the 254 observations that came from these patients, although not independent, raise the patient-to-variable ratio considerably. The fact that seven patients with missing data on the third occasion reduced the number of observations from 261 to 254 seems negligible. Second, patients of only two countries were studied, and third, only two forms of cancer were involved. Therefore, the feasibility of aggregating the

QLQ C-30 into five scales and/or two dimensions should be tested in future research with more patients, more countries, and more types of cancer.

IMPLICATIONS FOR CLINICAL PRACTICE

HRQOL measures, particularly the scales of the EORTC QLQ-C30, are frequently used as additional variables in clinical cancer research. However, they are seldom applied in the clinical interaction between patient and doctor. Although doctors are used to discuss laboratory values, imaging results, and medication with the patient, we envisage that also the patient's HRQOL is systematically reviewed during consultation. In such a case, having to inspect and discuss 15 pieces of information is time consuming and might ask too much of the patient's and clinician's attention. Five scores, or just two, are more manageable in the clinical context. That there exists a need for simple, quantitative indicators is witnessed by the frequent use of so called thermometers (i.e., for pain or distress).

QLQ-C30 summary scores could also play a role in the self-management of patients with any form of cancer. Nowadays it is not difficult to think of online or app-mediated administration of the QLQ-C30 that returns the relevant scores to the patient. It is a matter of debate whether this should be 15, five, or two scores, but fewer seem to be more manageable.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge all persons who have contributed to the design and data acquisition of this international research project, notably the authors of "Illness Perceptions and Quality of Life in Japanese and Dutch Patients with Non-Small-Cell Lung Cancer" (Kaptein et al., 2011) and "Illness Perceptions and Quality of Life in Japanese and Dutch Patients with Breast Cancer" (Kaptein et al., 2013).

REFERENCES

Aaronson, N., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., ... Takeda, F. (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*, 365–376.

Arraras, J. I., Arias, F., Tejedor, M., Pruja, E., Marcos, E., Martínez, E., & Valerdi, J. (2002). The EORTC QLQ-C30 (version 3.0) Quality of Life Questionnaire: Validation study for Spain with head and neck cancer patients. *Psycho-Oncology*, *11*, 249–256.

Arraras Urdaniz, J. I., Villafranca Iturre, E., Arias de la Vega, F., Dominguez Dominguez, M. A., Lainez Milagro, N., Manterola Burgaleta, A., ... Rojano y

Maite Martinez Aguillo, P. R. (2008). The EORTC Quality of Life Questionnaire QLQ-C30 (version 3.0). Validation study for Spanish prostate cancer patients. *Archivos Españoles de Urología*, *61*, 949–954.

Boehmer, S., & Luszczynska, A. (2006). Two kinds of items in quality of life instruments: 'Indicator and causal variables' in the EORTC QLQ-C30. *Quality of Life Research*, *15*, 131–141.

Broadbent, E., Petrie, K., Main, J., & Weinman, J. (2006). The Brief Illness Perception Questionnaire. *Journal of Psychosomatic Research*, *60*, 631–637.

Fayers, P. M., Aaronson, N. K., Bjordal, K., Groenvold, M., Curran, D., & Bottomley, A. (2001). *The EORTC QLQ-C30 scoring manual* (3rd ed.). Brussels, Belgium: European Organization for Research and Treatment of Cancer.

Fayers, P., Hand, D., Bjordal, K., & Groenvold, M. (1997). Causal indicators in quality of life research. *Quality of Life Research*, *6*, 393–406.

Ford, M., Havstadt, S., & Kart, C. (2001). Assessing the reliability of the EORTC QLQ-C30 in a sample of older African American and Caucasian adults. *Quality of Life Research*, *10*, 533–541.

Gotay, C. C., Blaine, D., Haynes, S. N., Holup, J., & Pagano, I. S. (2002). Assessment of quality of life in a multicultural cancer patient population. *Psychological Assessment*, *14*, 439–450.

Greenacre, M. (2010). *Biplots in practice*. Madrid, Spain: BBVA Foundation. Retrieved from http://www.multivariatestatistics.org/biplots.htm

Gundy, C. M., Fayers, P. M., Groenvold, M., Petersen, M. A., Scott, N. W., Sprangers, M. A. G., … Aaronson, N. K. (2012). Comparing higher order models for the EORTC QLQ-C30. *Quality of Life Research*, *21*, 1606–1617.

Henoch, I., Plone, A., & Tishelman, C. (2009). Increasing stringency in symptom cluster research: A methodological exploration of symptom clusters in patients with inoperable lung cancer. *Oncology Nursing Forum*, *36*, E283–E292.

Kaptein, A. A., Yamaoka, K., Snoei, L., Kobayashi, K., Uchida, Y., van der Kloot, W. A., … Rabe, K. (2011). Illness perceptions and quality of life in Japanese and Dutch patients with non-small-cell lung cancer. *Lung Cancer*, *72*, 384–390.

Kaptein, A. A., Yamaoka, K., Snoei, L., van der Kloot, W. A., Inoue, K., Kroep, J. R., … Nortier, H. (2013). Illness perceptions and quality of life in Japanese and Dutch women with breast cancer. *Journal of Psychosocial Oncology*, *31*, 83–102.

Karnofsky, D. A., Abelmann, W. H., Craver, L. F., & Burchenal, J. H. (1948). The use of the nitrogen mustards in the palliative treatment of carcinoma. *Cancer*, *1*, 634–656.

Kart, C., & Ford, M. (2002). Exploring the factorial structure of the EORTC QLQ-C30: Racial differences in measuring health-related quality of life in a sample of urban, older adults. *Journal of Aging and Health*, *14*, 399–421.

King, M. T., Dobson, A. J., & Harnett, P. R. (1996). A comparison of two quality-of-life questionnaires for cancer clinical trials: the Functional Living Index-Cancer (FLIC) and the Quality of Life Questionnaire Core Module (QLQ-C30). *Journal of Clinical Epidemiology*, *49*, 21–29.

Kobayashi, K., Morita, S., Shimonagayoshi, M., Kobayashi, M., Fujiki, Y., Uchida, Y., & Yamaguchi, K. (2008). Effects of socioeconomic factors and cancer survivors' worries on their quality of life (QOL) in Japan. *Psycho-Oncology*, *17*, 606–611.

Kobayashi, K., Takeda, F., Teramukai, S., Gotoh, I., Sakai, H., Yoneda, S., ... Yoshida, K. (1998). A cross-validation of the European Organization for Research and Treatment of Cancer QLQ-C30 (EORTC QLQ-C30) for Japanese with lung cancer. *European Journal of Cancer*, *34*, 810–815.

Linting, M., Meulman, J. J., Groenen, P. J. F., & van der Kooij, A. J. (2007). Nonlinear principal components analysis: Introduction and application. *Psychological Methods*, *12*, 336–358.

McCorkle, R., & Young, V. (1978). Development of a symptom distress scale. *Cancer Nursing*, *1*, 373–378.

McLachlan, S. A., Devins, G., & Goodwin, P. (1999). Factor analysis of the psychosocial items of the EORTC QLQ-C30 in metastatic breast cancer patients participating in a psychosocial intervention study. *Quality of Life Research*, *8*, 311–317.

Oerlemans, S., Mols, F., Nijziel, M. R., Lybeert, M., & van de Poll-Franse, L. V. (2011). The impact of treatment, socio-demographic and clinical characteristics on health-related quality of life among Hodgkin's and non-Hodgkin's lymphoma survivors: A systematic review. *Annals of Hematology*, *90*, 993–1004.

Osaba, D., Zee, B., Pater, J., Warr, D., Kaizer, L., & Latreille, J. (1994). Psychometric properties and responsiveness of the EORTC Quality of Life Questionnaire (QLQ-C30) in patients with breast, ovarian and lung cancer. *Quality of Life Research*, *3*, 353–364.

Pagano, I. S., & Gotay, C. C. (2006). Modeling quality of life in cancer patients as a unidimensional construct. *Hawai'i Medical Journal*, *65*, 74–82.

Ringdal, G., & Ringdal, K. (1993). Testing the EORTC quality of life questionnaire on cancer patients with heterogeneous diagnoses. *Quality of Life Research*, *2*, 129–140.

Schipper, H., Clinch, J., McMurray, A., & Levitt, M. (1984). Measuring the quality of life of cancer patients: The Functional Living Index-Cancer: Development and validation. *Journal of Clinical Oncology*, *2*, 472–483.

Strasser, F., Müller-Käser, I., & Dietrich, D. (2009). Evaluating cognitive, emotional, and physical fatigue domains in daily practice by single-item questions in patients with advanced cancer: A cross-sectional pragmatic study. *Journal of Pain and Symptom Management*, *38*, 505–514.

Van Steen, K., Curran, D., Kramer, J., Molenberghs, G., van Vrecken, A., Bottomley, A., & Sylvester, R. (2002). Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. *Statistics in Medicine*, *21*, 3865–3884.

APPENDIX

Construct Validity Studies: Recovering the QLQ-C30 Scales from the Items

Ford et al. (2001) factor analyzed the data of 255 African American and 234 Caucasian non-cancer patients. They confirmed a three-factor model for the 11 items of PF, EF, and FA. The remaining items were not included in the analysis. The article contains no information about the interrelatedness of PF, EF, and FA.

Kart and Ford (2002) used the same patients of Ford et al. (2001). They ran separate principal component analyses in the two samples. They obtained seven principal components (factors) for the 30 items of the QLQ-C30. In the Caucasian sample, one factor replicated EF and a second factor consisted of FA and four items of PF. In the African American group, EF and CF were replicated as separate factors. In both groups, the remaining factors show mixtures of the items of various scales and symptoms.

Gotay et al. (2002) conducted three studies on the content validity of the QLQ-C30. Study 2 used confirmatory factor analysis on the data of 367 heterogeneous cancer patients in Hawaii. One latent factor was found that could explain the relations between the five functioning scales, QL, and a variable that counted the number of symptoms endorsed by a patient. By combining PF and RF into one scale, the model was improved and showed excellent fit.

In a study of 150 women with metastatic breast cancer McLachlan et al. (1999) analyzed the intercorrelations among the 12 items of EF, CF, RF, SF, and QL. They found two principal components that were labeled "emotional distress" {EF, CF} and "functional ability" {RF, SF, QL}.

Using the data before and after chemotherapy of 535 patients with several types of cancer, Osaba et al. (1994) performed several factor analyses on the items of the nine multi-item scales of the QLQ-C30. They reported nine orthogonal factors that reproduced the nine postulated scales reasonably well. For the total group of patients, the matrix of intercorrelations among the 15 QLQ-C30 scales showed high correlations ($|r| > .60$) between PF, RF, and FA and between FA and QL.

Studies of the Interrelations among the QLQ-C30 Scale

Boehmer and Luszczynska (2006) tested the model of Fayers, Hand, Bjordal, and Groenvold (1997) who proposed that the functioning scales of the QLQ-C30 are indicators of the patient's quality of life, whereas the symptom scores are seen as causal variables that influence quality of life. Boehmer and Luszcynska tested two structural equation models. The first model hypothesized that there is one latent factor "symptomatology" underlying nine latent variables for the symptom scales, which are supposed to influence one latent construct HRQOL. This construct in its turn effects separate latent factors for PF, RF, CF, SF, and EF, which are measured by the corresponding items. The second model hypothesized that both the functioning and the symptom scales can be explained by one underlying latent construct. The first model was supported by confirmatory factor analyses, although it was not significantly better than the simpler model in which one latent factor underlies all scales and items of the QLQ-C30.

Study 3 of Gotay et al. (2002) used the same 367 heterogeneous cancer patients as Study 2. Factor analysis of QL, CF, EF, SF, two combinations of PF and RF items, and 22 new questions, yielded one single factor with eigenvalue >1, which demonstrates the connectedness of the six QLQ-C30 scales.

Gundy et al. (2012) used the data of 4541 cancer patients who were heterogeneous with respect to cancer type and nationality to compare seven structural equation models that differed with regard to the postulated relations between 14 QLQ-C30 scales (FI was excluded) and the number and structure of explanatory factors. The best fitting model, labeled the Physical/Mental health model, contained three correlated higher-order factors: Physical Health, Mental Health, and QL. PF, NV, DY, AP, CO, and DI were explained by Physical Health; CF and EF were explained by Mental Health. To explain RF, SF, FA, PA, and SL both factors were needed. The estimated correlation between Physical Health and Mental Health was .74; the correlations of those factors with QL were not reported.

In a sample of 177 heterogeneous Norwegian cancer patients, Ringdal and Ringdal (1993) used Mokken Scale Analysis on subsets of the 30 items. They replicated the CF, EF, SF, QL, PA, and FA scales and found evidence that PF and RF belonged to one "personal functioning scale" and that NV and AP could be combined. Principal component analysis of the above scales except QL, yielded two oblique factors ($r = .47$) with high loadings of {PF & RF}, {NV & AP}, PA, and FA on the first factor, and high loadings of FA, CF, EF, and SF on the second factor. QL was related to both factors. The authors interpreted the first factor as a physical one and the second factor as psychological. They regarded the positive correlation of these factors "as a weak argument for the existence of a general QOL dimension" (Ringdal & Ringdal, 1993, p. 139).

Using data from 187 women with advanced breast cancer, van Steen et al. (2002) found three principal components in 12 scales of the QLQ-C30 (CO, DI, and FI were not studied). The first factor contained high absolute loadings of PF, RF, SF, QL, and PA. The second factor showed a cluster of AP, DY, FA, NV and CF. The third factor primarily consisted of SL. EF had substantial loadings on Factor 2 and Factor 3.

Studies Relating QLQ-C30 Scales and Items with Other Instruments

In a study of 201 head and neck cancer patients, Arraras et al. (2002) reported correlations of .67 and .70 between PF and RF on two subsequent occasions. The correlations between PF and FA were −.49 and −.67, between RF and SF .49 and .70, and between RF and FA −.40 and −.68.

A study of 137 prostate cancer patients by Arraras Urdaniz et al. (2008) yielded correlations between PF and RF of .65 on two subsequent occasions.

The correlations between PF and SF were .57 and .50. PF and FA correlated −.71 and −.63, RF and FA −.65 and −.62, and EF and FA −.51 and −.57.

Henoch et al. (2009) analyzed the correlations among those items of the QLQ-C30 that were also present in the Symptom Distress Scale (McCorkle & Young, 1978). They obtained three clusters of items. The first cluster—"mood"—contained two items of the EF scale, one item of the CF scale, and SL. The second cluster—"pain"—contained FA, AP, one of the PA items, one of the NV items, and the mean score of the DI and CO items. The third cluster—"respiratory"—consisted of DY and the non-QLQ-C30 item "cough".

King et al. (1996) performed a multitrait-multimethod study on the scales of the QLQ-C30 and seven subscores of the Function Living Index-Cancer (FLIC; Schipper, Clinch, McMurray, & Levitt, 1984). The correlations among the QLQ-C30 scales were substantial and suggested the existence of three clusters {RF, PF}, {EF, SF}, and {QL, FA, PA}, which possibly could be merged into one cluster.

Using a structural equation model for the effects of socioeconomic factors on the HRQOL of 130 Japanese cancer survivors, Kobayashi et al. (2008) presented evidence for one latent HRQOL factor underlying PF, RF, EF, CF, and SF.

Pagano and Gotay (2006) applied item response theory to the data of 366 heterogeneous cancer patients in Hawaii to obtain a unidimensional scale from the items of the QLQ-C30 and 35 items of two other instruments. The resulting 22-item scale contained 15 items from the QLQ-C30: the two QL items, one PF items, one RF items, both SF items, all four EF items, all three FA items, one PA items, and the item for AP.

Strasser et al. (2009) used PF, EF, and CF scores of 61 patients with advanced cancer in a principal component analysis with several scales of other measuring instruments, and obtained two factors—"cognitive" and "physical"—with high loadings of CF and PF, respectively. EF had a high loading (.71) on the first factor and a substantial loading (.49) on the second factor.