Statistical analysis

If the OS of the patients treated with ACT was significantly longer than that of the patients treated with TAM, ACT would be recommended as the new standard treatment. The estimated 5-year OS of these patients is commonly 64–88 % [7–9]. The initial sample size was calculated as 280 patients to detect a prolongation of 5-year OS from 75 % in the TAM arm to 87 % in the ACT arm with 80 % power and a two-sided alpha of 5 %. The planned study period was originally 2 years for accrual and an additional 5 years for follow-up. Due to the slow accrual, the protocol was revised to prolong the accrual period, and the sample size was revised to 220 patients with an accrual period of 5 years. OS and RFS were estimated using Kaplan–Meier method, and curves were compared by using a log-rank test. Hazard ratios of treatment effects were estimated through a Cox regression model. All analyses were based on intention to treat. All statistical analyses were performed using the SAS software package, release 8.2.

Interim analysis and monitoring

It was planned that an interim analysis would be performed when half of the total number of patients had been enrolled. The Data and Safety Monitoring Committee (DSMC) of the JCOG independently reviewed the interim analysis report, and premature termination of the trial was considered at that stage. In-house interim monitoring was performed by the Data Center to ensure data submission, patient eligibility, protocol compliance, safety, and on-schedule study progress. The monitoring reports were submitted to and reviewed by the DSMC every 6 months.

Results

Patient population

This study was initially implemented in 1994. After approximately 110 patients had been enrolled, the protocol was revised to prolong patient recruitment until December 1999. At the interim analysis on June 1999, patient accrual was so slow that DSMC recommended that patient accrual should be terminated or continued with the primary endpoint changed to RFS. Furthermore, at a consensus meeting in St. Gallen in 1997, it was established that the administration of TAM to hormone receptor-negative patients was ethically unacceptable. Therefore, the recruitment of patients was terminated based on suggestions from the DSMC of JCOG.

In total, 131 patients were recruited. Two patients were registered twice. Thus, 129 patients were randomized

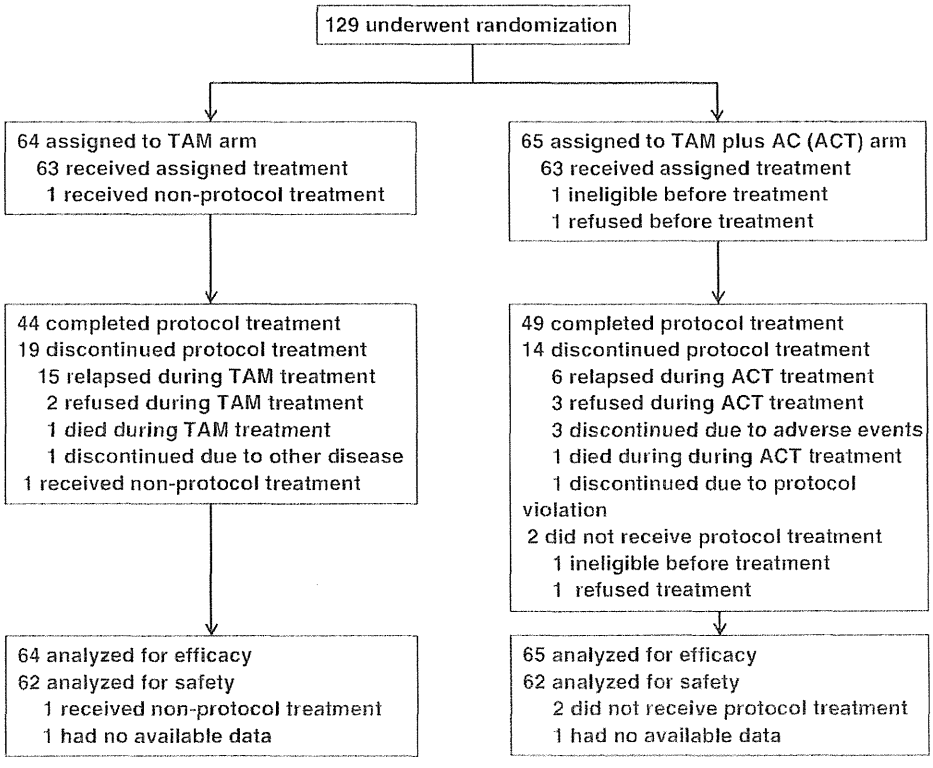

Fig. 1 Trial profile of the Japan Clinical Oncology Group study JCOG 9401

Table 1 Patient characteristics

	TAM ($n = 64$)	ACT ($n = 65$)
Age (years)		
Median	58	59
Range	47–70	46–70
No. of positive axillary nodes		
1–3	46	46
4–9	18	19
ER and/or PgR		
Negative/unknown	18	19
Positive	46	46
HER2		
Negative/unknown	53	57
Positive	11	8
Stage		
I	10	12
II	44	43
IIIA	10	10
Operation		
Radical mastectomy	3	6
Total mastectomy	59	55
Partial resection	2	4

(Fig. 1). One patient was ineligible because of the previous administration of TAM, but was still included in the analysis. The baseline characteristics were well balanced between the two groups (Table 1). The median age was 59 years (46–70 years). The number of patients with node metastases involving 1–3 nodes was 92 (71.3 %) and the number with 4–9 involved nodes was 37 (28.7 %). The number of patients with both ER− and PR− tumors, including patients with unknown hormone status, was 37 (28.7 %). Most patients (95.3 %) underwent total mastectomy. Sixty-four cases were assigned to the ACT arm and 65 cases were assigned to the TAM arm. The data from an immunohistochemistry assay of HER2 protein were missing for 2 cases in the ACT arm and for 1 case in the TAM arm; the data from a cytosol assay of HER2 protein were missing for 2 cases in each group. However, this did not influence the results of this study.

Treatment completion

The protocol treatment in the TAM arm was completed in 44 of the 64 cases (68.8 %). The protocol treatment in the ACT arm was completed in 49 of the 65 cases (75.4 %) (Fig. 1).

Survival

Sixty-four and 65 patients were enrolled and analyzed in the TAM and ACT arms, respectively, with no significant

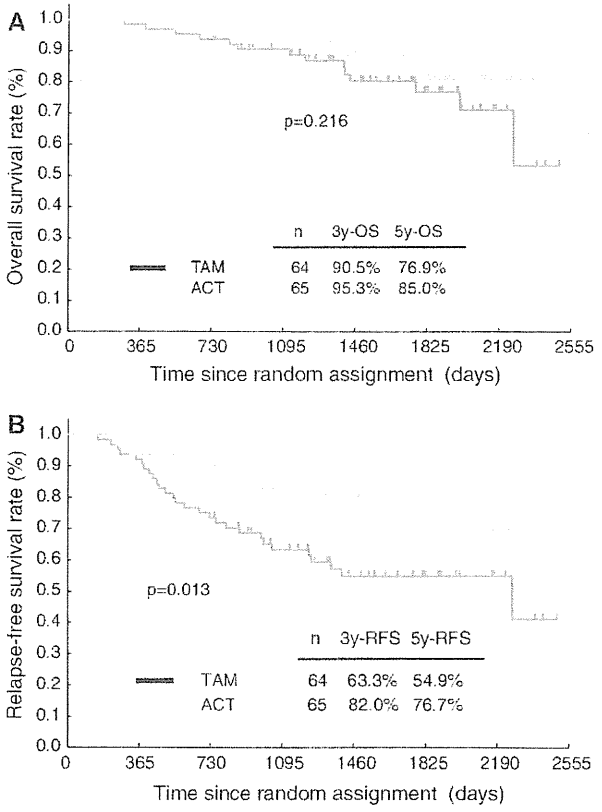

Fig. 2 Kaplan–Meier curves of overall survival (**a**) and relapse-free survival (**b**) for node-positive breast cancer patients treated with tamoxifen or tamoxifen with anthracycline and cyclophosphamide

difference in overall survival between them [$p = 0.216$, hazard ratio 0.58, 95 % confidence interval (CI) 0.24–1.39] (Fig. 2a). The 3- and 5-year OS were 90.5 and 76.9 % for the TAM arm and 95.3 and 85.0 % for the ACT arm, respectively. The RFS in the ACT arm was significantly longer than that in the TAM arm ($p = 0.013$, hazard ratio 0.45, 95 % CI 0.24–0.86) (Fig. 2b). The 3- and 5-year RFS were 63.3 and 54.9 % in the TAM arm and 82.0 and 76.7 % in the ACT arm, respectively.

Subgroup analysis was performed according to hormone receptor status. The numbers of ER+ and/or PR+ patients in the TAM and ACT arms were 46 (71.9 %) and 46 (70.8 %), respectively. The OS is shown in Fig. 3a. ER− patients included in the TAM arm had a worse prognosis than those in the other three groups. ER− patients on ACT showed a better overall survival than those on TAM alone. The OSs of the ER+ and/or PR+ groups were good, regardless of the arm considered. RFS in the ER+ and/or PR+ patients in the TAM arm was worse than those in the ACT arm (Fig. 3b). The ER− and PR− patients included in the TAM arm had the worst prognosis.

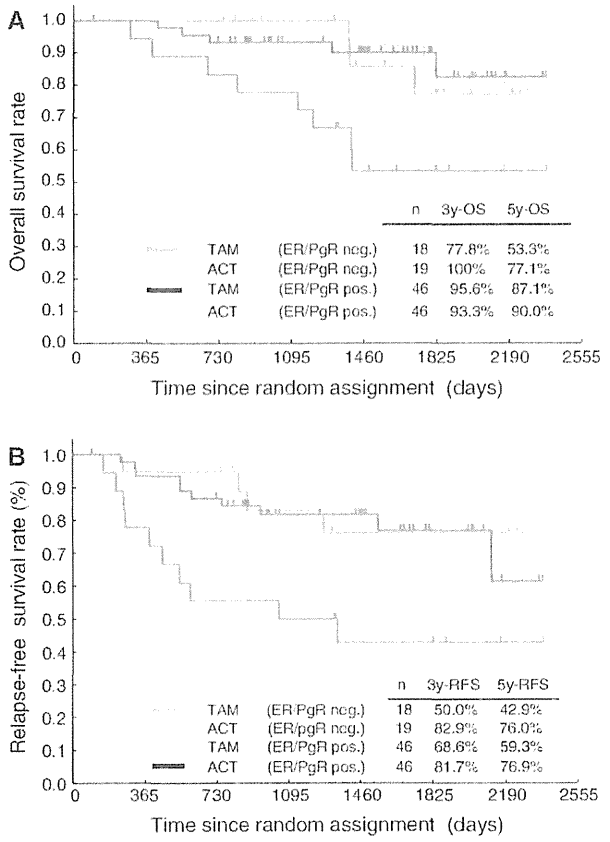

Fig. 3 Kaplan–Meier curves of overall survival (**a**) and relapse-free survival (**b**) for node-positive breast cancer patients treated with tamoxifen or tamoxifen with anthracycline and cyclophosphamide, according to estrogen receptor (ER) and progesteron receptor (PgR) status

Table 2 Hematological (A) and nonhematological (B) toxicities

Toxicities	Grade 2 (%)	Grade 3 (%)	Grade 4 (%)
(A) Hematological toxicities ($n = 62$)			
TAM			
WBC	0 (0)	0 (0)	0 (0)
Hb	1 (2)	0 (0)	0 (0)
T-Bil	2 (3)	0 (0)	0 (0)
GOT	3 (5)	0 (0)	0 (0)
GPT	4 (6)	0 (0)	0 (0)
ACT			
WBC	9 (15)	3 (5)	0 (0)
Hb	2 (3)	0 (0)	–
T-Bil	7 (11)	0 (0)	0 (0)
GOT	7 (11)	0 (0)	0 (0)
GPT	7 (11)	0 (0)	0 (0)
(B) Non-hematological toxicities ($n = 62$)			
TAM			
Infection	0 (0)	0 (0)	0 (0)
Nausea/vomiting	0 (0)	0 (0)	–
Thrombosis	0 (0)	0 (0)	0 (0)
Alopecia	0 (0)	–	–
ACT			
Infection	1 (2)	0 (0)	0 (0)
Nausea/vomiting	20 (32)	1 (2)	–
Thrombosis	1 (2)	0 (0)	0 (0)
Alopecia	29 (42)	–	–

Safety

Toxicities detected are listed in Table 2. No grade 4 events were noted. Higher proportions of the patients in the ACT arm experienced grade 3 decreased white blood cell count (TAM 0 %, ACT 4.8 %), grade 2–3 nausea (TAM 0.0 %, ACT 33.9 %), and grade 2 alopecia (TAM 0.0 %, ACT 46.8 %) than the corresponding patients in the TAM arm. Higher proportions of the patients in the TAM arm had grade 2 increased GOT and GPT (TAM 11.3 and 11.3 %, ACT 4.8 and 6.5 %).

Discussion

It is uncertain whether adjuvant chemotherapy is required in the treatment of postmenopausal breast cancer with hormone-responsive and intermediate risk. There have been few clinical trials to compare hormone therapy alone with chemotherapy plus hormone therapy [10, 11]. These studies suggested that the efficacy of chemoendocrine therapy as an adjuvant therapy for improving the prognosis of highly ER-positive patients was limited [11]. A meta-analysis of metastatic breast cancer cases indicated that chemoendocrine therapy did not lead to an improved prognosis compared to monotherapy [12]. We reanalyzed and reported this old study starting from 1994, because we thought that this would provide a valuable source of information when attempting to answer this clinical question and determine the optimal adjuvant treatment strategy for those patients. This study was designed to demonstrate the superiority of our anthracycline-containing regimen, ACT, over TAM alone (regardless of ER and PR status), which was the most common adjuvant treatment for postmenopausal patients in the early 1990s. The planned recruitment for each study arm was 110 patients at the beginning of the study and patient recruitment was started in December 1994. This trial was terminated in July 1999 because patient accrual was slow and TAM was contraindicated for ER− PMBC during the course of this trial [13]. The NSABP B-23 reported no improvement in the prognosis of ER− patients receiving adjuvant TAM. A meta-analysis conducted by the EBCTCG showed that the risk

reduction for recurrence on using TAM in ER− patients was 6 %; however, these results were biased by the error associated with ER- and PR-like enzyme immunoassay methods [14]. After those reports, TAM was not used for ER− patents. In this study, ER− patients (28.7 %) who received only TAM had the worst prognosis, as shown in the subgroup analysis. These patients may include both completely negative ER and 1–9 % positive patients for whom adjuvant hormone therapy has recently been indicated. The prognosis for the patients with completely negative ER was probably similar to that of the patients who did not receive adjuvant treatment, and the efficacy of TAM for the patients with slightly positive ER was unclear based on these results.

Overall, the analysis showed that there was no significant prognostic effect of adjuvant chemotherapy on OS in spite of the inclusion of ER− patients. The effect on the prognosis was greater for ER− patients than for ER+ patients, and TAM is not effective for ER− patients. Currently, the common duration of TAM treatment for ER+ patients is 5 years. Two years of adjuvant TAM might not have maximized the prognosis for these patients. However, these factors could not have influenced the superiority of additional chemotherapy in this study. Rather, the inclusion of ER− patients and the insufficient number of enrolled patients may have influenced this result. In addition, we believe that the insignificant effect of chemotherapy in ER+ rather than in ER− patients is important.

The effect of adjuvant chemotherapy on the prognosis for ER+ patients may be less than that for ER− patients, and the additional effect on top of that of endocrine therapy is relatively small. A meta-analysis by the EBCTCG showed that the reduction in the risk for recurrence from the use of adjuvant chemotherapy in postmenopausal ER− breast cancer patients was twice that in ER+ patients [2]. Based on the drug effects predicted from tumor biology, the new treatment strategy was recommended at the St. Gallen consensus meeting of 2011 [15]. ER+ patients have a low Ki67 labeling index, which is correlated with the efficacy of chemotherapy, and HER2-negative breast cancer patients have a low sensitivity to chemotherapy and a high sensitivity to endocrine therapy. In this study, no significant effect on OS was noted upon the addition of chemotherapy to endocrine therapy in node-positive patients with a high risk for recurrence. The ER− subgroup showed improved prognosis with additional chemotherapy, but the ER+ subgroup did not. The chemotherapy dose was lower than recently recommended, because the standard AC regimen was not established at the beginning of this study in Japan. Thus, the maximum effect of the adjuvant chemotherapy was not fully elicited in this study. However, we believe that one of the reasons for this negative result is that some ER+ subgroups can be treated adequately with endocrine therapy. In addition, we need extra information to judge the indication for adding chemotherapy to endocrine therapy in ER+ breast cancer patients. The degree of ER expression is one of them. It is reported that the efficacy of adjuvant chemotherapy in highly ER+ breast cancer cases is limited [11]. The most promising methods to use to identify these subgroups are the multigene assay and the Ki67 labeling index. Oncotype DX, which is a multigene assay to predict the efficacy of adjuvant chemotherapy against ER+ breast cancer, has been evaluated in prospective studies in which it showed demonstrable utility [16, 17]. However, these assays are not commonly performed because of the cost and inconvenience involved. The Ki67 labeling index is reported to be a prognostic factor. Moreover, it may be a predictive factor for chemotherapy [18, 19] and is especially meaningful in ER+ patients [20]. However, the limitations of this method include a lack of measurement clarity and thresholds [21]. Our study was started in 1994, and the surgical specimens were very old, so we cannot reanalyze these factors. We need to plan a study to establish a strategy for appropriate adjuvant treatment of ER+ patients using these new indices.

TAM alone was less toxic than ACT. No grade 3 toxicities were noted in patients treated with TAM alone, and the frequency of all toxicities was less in the TAM arm than in the ACT arm. A previous study reported an increased rate of thromboembolic complications with chemoendocrine combination adjuvant therapy [10]. There was only one G2 thrombosis in the patients with ACT.

In this study, there was no significant improvement in PMBC patient prognosis upon adding chemotherapy. Both benefits and risks need to be considered when choosing whether to implement adjuvant treatment. Many chemotherapy-associated toxicities are harmful and sometimes fatal. In the absence of an effect on the prognosis to offset the risks associated with additional chemotherapy, adjuvant chemotherapy should not be administered to all ER+ patients. More detailed analysis and definitive prospective trials are warranted to validate our findings.

Acknowledgments We thank Ms. Kyoko Minamoto and Kazumi Kubota for data management, Dr. Naoki Ishizuka and Mr. Junki Mizusawa for statistical analyses, and Dr. Kenichi Nakamura for the preparation of the manuscript. This study was supported by a National Cancer Center Research and Development Fund (23-A-16 and 23-A-17) and Grants-in-Aid for Cancer Research (5S-1, 8S-1, 11S-1, 11S-4, 14S-1, 14S-4, 17S-1, 17S-5, 20S-1 and 20S-6) from the Ministry of Health, Labour and Welfare, Japan.

Conflict of interest Hiroji Iwata received honoraria for speaking events from Chugai Pharmaceutical Co., Ltd. Tadahiko Shien, Kenjiro Aogi, Takashi Fukutomi, Kenichi Inoue, Takayuki Kinoshita, Masato Takahashi, Akira Matsui, Taro Shibata, Haruhiko Fukuda had no conflicts of interest.

Appendix: Participating institutions (from north to south)

The 22 institutions that belonged to the JCOG Breast Cancer Study Group and participated in this trial are as follows: National Sapporo Hospital, International Medical Center of Japan, Tochigi Cancer Center, Metropolitan Komagome Hospital, National Cancer Center, National Cancer Center East, Tokai University Hospital, National Atami Hospital, Hamamatsu Medical Center, Aichi Cancer Center, Osaka National Hospital, Kinki University Hospital, National Shikoku Cancer Center, National Kure Medical Center, National Nagasaki Medical Center, Saitama Cancer Center, St Luke's International Hospital, Hyogo Medical Center, Shizuoka Cancer Center, Niigata Cancer Center Hospital, Kawasaki Medical School Hospital, and Kitakyushu Municipal Medical Center.

References

1. Early Breast Cancer Trialists' Collaborative Group (1990) Treatment of early breast cancer: worldwide evidence in 1985–1990. A systematic overview of all available randomized trials in early breast cancer of adjuvant endocrine and cytotoxic therapy-treatment of early breast cancer, vol 1. Oxford University Press, Oxford, pp 21–49
2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365(9472):1687–1717
3. Consensus conference (1985) Adjuvant chemotherapy for breast cancer. JAMA 254:3461–3463
4. Glick JH, Gelber RD, Goldhirsch A et al (1992) Meeting highlights: adjuvant therapy for primary breast cancer. J Natl Cancer Inst 84:1479–1485
5. Fisher B, Redmond C, Legault-Poisson S et al (1990) Postoperative chemotherapy and tamoxifen compared with tamoxifen alone in the treatment of positive node breast cancer patients aged 50 years and older with tumors responsive to tamoxifen: results from the national surgical adjuvant breast and bowel project B-16. J Clin Oncol 8:1005–1018
6. Tobinai K, Kohno A, Shimada Y et al (1993) Toxicity grading criteria of the Japan Clinical Oncology Group. The Clinical Trial Review Committee of the Japan Clinical Oncology Group. Jpn J Clin Oncol 23(4):250–257
7. Gelber RD, Goldhirsch A (1986) A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. J Clin Oncol 4:1772–1779
8. Boccardo F, Rubagotti A, Amoroso D et al (1992) Chemotherapy versus tamoxifen versus chemotherapy plus tamoxifen in node positive, oestrogen receptor positive breast cancer patients. An update at 7 years of the 1st GROCTA Trial. Eur J Cancer 28:673–680
9. Pearson OH, Hubay CA, Gordon NH et al (1989) Endocrine versus endocrine plus five-drug chemotherapy in postmenopausal women with stage II estrogen receptor positive breast cancer. Cancer 64:1819–1823
10. Pritchard KI, Paterson AH, Paul NA et al (1996) Increased thromboembolic complications with concurrent tamoxifen and chemotherapy in a randomized trial of adjuvant therapy for women with breast cancer. National Cancer Institute of Canada Clinical Trials Group Breast Cancer Site Group. J Clin Oncol 14(10):2731–2737
11. Pagani O, Gelber S, Simoncini E et al (2009) Is adjuvant chemotherapy of benefit for postmenopausal women who receive endocrine treatment for highly endocrine-responsive, node-positive breast cancer? International Breast Cancer Study Group Trials VII and 12-93. Breast Cancer Res Treat 116(3):491–500
12. Fossati R, Confalonieri C, Torri V et al (1998) Cytotoxic and hormonal treatment for metastatic breast cancer: a systematic review of published randomized trials involving 31,510 women. J Clin Oncol 16(10):3439–3460
13. Fisher B, Anderson S, Tan-Chiu E et al (2001) Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. J Clin Oncol 19(4):931–942
14. Early Breast Cancer Trialists' Collaborative Group (1998) Tamoxifen for early breast cancer: an overview of the randomized trials. Lancet 351(9114):1451–1467
15. Goldhirsch A, Wood WC, Coates AS et al (2011) Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol 22(8):1736–1747
16. Paik S, Tang G, Shak S et al (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24(23):3726–3734
17. Albain KS, Barlow WE, Shak S et al (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 11(1):55–65
18. Hugh J, Hanson J, Cheang MC et al (2009) Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. J Clin Oncol 27:1168–1176
19. Penault-Llorca F, André F, Sagan C et al (2009) Ki-67 expression and docetaxel efficacy in patients with estrogen receptor-positive breast cancer. J Clin Oncol 27(17):2809–2815
20. Niikura N, Iwamoto T, Masuda S et al (2012) Immunohistochemical Ki67 labeling index has similar proliferation predictive power to various gene signatures in breast cancer. Cancer Sci 103(8):1508–1512
21. Dowsett M, Nielsen TO, A'Hern R et al (2011) Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer Working Group. JNCI. J Natl Cancer Inst 103:1656–1664

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

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

Routledge
Taylor & Francis Group

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 inves-

tigated 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**	.870[a]**	.462**	−.623**
RF	.571**	.646**	.955[a]**	.507**	−.638**
SF	.577**	1	.631**	.549**	−.685**
CF	.464**	.533**	.590**	.906[a]**	−.637**
EF	.422**	.453**	.359**	.894[a]**	−.516**
FA	−.638**	−.601**	−.756**	−.578**	.824[a]**
NV	−.339**	−.369**	−.284**	−.372**	.580[a]**
PA	−.570**	−.517**	−.654**	−.577**	.748[a]**
DY	−.353**	−.186**	−.413**	−.199*	.450[a]**
SL	−.470**	−.469**	−.440**	−.530**	.712[a]**
AP	−.575**	−.448**	−.535**	−.458**	.753[a]**
CO	−.235**	−.254**	−.246**	−.266**	.454[a]**
DI	−.139	−.119	−.161	−.121	.272[a]**
FI	−.323**	−.367**	−.127	−.263**	.461[a]**
			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

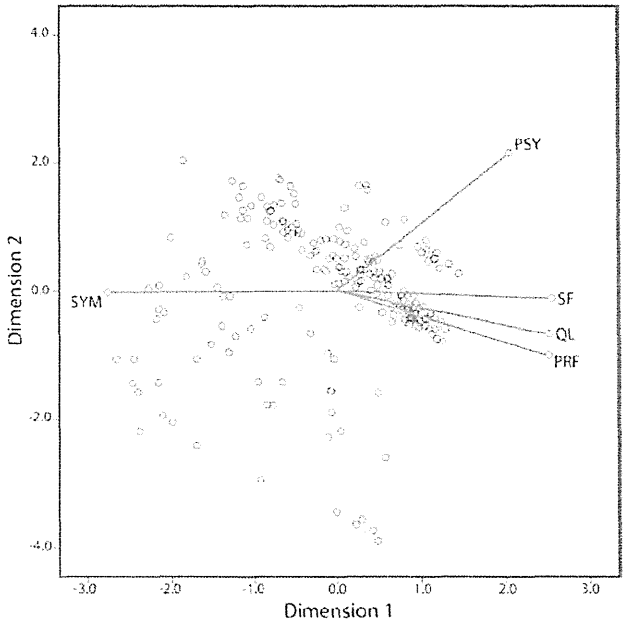

FIGURE 1 Biplot of patients and EORTC QLQ-C30 summary variable in a two-dimensional space. QL = global health/quality of life; SF = social functioning; PRF = physical and role functioning; PSY = psychological functioning; SYM = symtomatology.

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

Variable	Standardized Regression Coefficients		Multiple Correlation Coefficient
	Component 1	Component 2	
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