

Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting?

A. M. Stiggelbout · E. de Vogel-Voogt ·
E. M. Noordijk · T. P. M. Vliet Vlieland

Accepted: 3 March 2008 / Published online: 9 April 2008
© The Author(s) 2008

Abstract In the schedule for the evaluation of individual quality of life (SEIQoL) the weights for five individualized quality of life domains have been derived by judgment analysis and direct weighting (DW). We studied the feasibility and validity of adaptive conjoint analysis (ACA) as an alternative method to derive weights in 27 cancer patients and 20 patients with rheumatoid arthritis. Further, we assessed the convergence between direct weights and weights derived by ACA, and their correlation with global quality-of-life scores. All respondents finished the ACA task, but one in five respondents were upset about the ACA task. Further, the task was vulnerable to judgment ‘errors’, such as inconsistent answers. The agreement between the two weights was low. Both weighted index scores were strongly correlated to the unweighted index score. The relationships between the index score and scores on a visual analogue scale for global individual quality of life and global quality of life were similar whether or not the index score was calculated with DW weights, with ACA weights, or without using weights. We conclude that, because weights did not improve the correlation between the index score and global quality of life scores, it seems

sufficient to use the unweighted index score as a measure for global individual quality of life.

Keywords Conjoint analysis · Direct weighting · Individual quality of life · SEIQoL · Patients

Abbreviations

ACA	Adaptive conjoint analysis
DW	Direct weighting
iQoL	Individual quality of life
JA	Judgment analysis
QoL	Quality of Life
SEIQoL	Schedule for the evaluation of individual quality of Life
SEIQoL-DW	Schedule for the evaluation of individual quality of life-direct weighting
VAS	Visual analogue scale

Introduction

Quality of life (QoL) instruments have traditionally been based on a needs model of QoL, with the same domains, assessment criteria, and weighting applied to all respondents [1]. Individual quality of life (iQoL) measures recognize that people define life domains in different ways, use different criteria to evaluate the domains, and place differing emphasis on their importance to overall quality of life [2, 3]. The schedule for the evaluation of individual quality of life (SEIQoL) is a method of assessing iQoL. It has been used in many contexts [3–7].

The SEIQoL consists of three stages of administration. The first is a semistructured interview, in which respondents are asked to nominate five areas of life (domains) that they consider most important to the overall quality of their

A. M. Stiggelbout (✉) · E. de Vogel-Voogt
Department of Medical Decision Making, J10-S, Leiden
University Medical Center, P.O. Box 9600, 2300 RC Leiden,
The Netherlands
e-mail: a.m.stiggelbout@lumc.nl

E. M. Noordijk
Department of Clinical Oncology, Leiden University Medical
Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands

T. P. M. Vliet Vlieland
Department of Rheumatology, Leiden University Medical
Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands

lives [1]. In the second stage the respondents rate their functioning on each of these domains, and in the third the relative weights of the domains are quantified.

In the original version of the SEIQoL a method called judgment analysis (JA), based on social judgment theory, was used for stage 3 [2, 3]. In social judgment theory, linear models are used to explain the impact of important factors, or cues, and their weights, on judgments. For the SEIQoL, to quantify the relative weights, subjects were presented with 30 randomly generated profiles of hypothetical situations labeled with the five chosen domains [3]. Respondents were asked to rate the QoL they associated with each profile on a VAS. Ten of the 30 scenarios were replicates to allow for a measure of consistency of judgments. The judgments were modeled using multiple regression to produce five relative weights summing to 100 [2]. The authors stated that JA was a successful aid to the evaluation of QoL [2, 3, 8], but from a theoretical viewpoint the number of 20 scenarios was small. Thirty to 50 scenarios are recommended to model the five weights [9]. McGee reported the use of 40 scenarios in the first study on the SEIQoL [8], but in later studies only 30 scenarios were used, of which ten were replicates [1–3]. Further, the time required to administer the JA to respondents is long and the task cumbersome, especially for older and cognitively impaired people. It requires compensatory decision-making processes, i.e., the ability to make an overall judgment on the basis of weighted information. These drawbacks make JA unsuitable for application in most clinical practices as well as for frequent assessment over a short period of time [1]. Because of these problems with JA, Browne et al. [1] developed a more easy to administer version of the SEIQoL, namely SEIQoL-direct weighting (SEIQoL-DW). Respondents are asked to fill in a pie chart in which the relative size of each sector of the pie represents the weight the respondent attaches to a QoL domain. The validity and reliability of the DW in comparison with the JA-based SEIQoL have only been investigated in two small studies. The first was done in a sample of 40 healthy volunteers [1]. The mean absolute difference between the 200 weights derived for each method, that is, 40 times five weights, was 7.8 at the first measurement, and 7.2 at the second measurement. When weights derived from the two methods were converted to ranks and compared for agreement, the κ value was moderate at both measurements (0.40 and 0.44, respectively). The reliability of the weights was moderate for the DW ($\kappa = 0.51$), and only fair for JA ($\kappa = 0.31$) [1]. In a later study, Waldron et al. compared the psychometric characteristics of the SEIQoL-DW with the SEIQoL-JA among 80 patients with advanced incurable cancer. The index scores generated by the two methods fell within a range of 14.9 [10]. These differences are

large in clinical terms [3]. Therefore, the authors [10] concluded that the two methods are not interchangeable. Despite this lack of data on the DW, and its demonstrated lack of agreement with the JA, it has become the standard method for eliciting weights of the SEIQoL, due to it being simple to administer. However, a fundamental distinction in cognitive psychology is that between explicit and implicit thought [1], and previous evidence with weighting methods suggests that respondents may be unable to provide accurate implicit weights through a method such as the DW [1]. Indirect methods such as JA are based on more basic and simple judgmental tasks, such as paired comparisons, and thereby may reduce possible biases that may play a role in direct judgments. Further, indirect methods are more likely to avoid the social desirability effect according to which respondents bias their response toward the perceived values of the researcher/clinician [11]. Another advantage is that indirect methods can provide measures of internal reliability and validity for individual interviews [1]. An alternative indirect weighting method might be conjoint analysis, which was developed in mathematical psychology and, like JA, has a strong theoretical basis [12–14]. As in JA, conjoint analysis is based on the premises that any treatment or health state can be described by its characteristics (or attributes) and that the extent to which an individual values a treatment or health state depends on the levels of these characteristics. The method can be used to estimate the relative importance of these attributes, and may therefore be suitable for eliciting the weights of domains of iQoL, since the domains can be seen as attributes of iQoL.

In conjoint analysis the number of paired comparisons needed to estimate the weights may be as high as the number of scenarios needed in JA, but nowadays software, adaptive conjoint analysis, is available that is adaptive to respondent's answers and can minimize the number of paired comparisons. As an indirect method, it may provide a feasible alternative to JA. ACA has been used to derive treatment preferences in patients with lupus nephritis, with HIV medication, and with cancer [15–17].

This study aims to assess the feasibility and the validity of the ACA to derive weights for iQoL domains. Furthermore, agreement of the weighting procedures performed by the ACA and the DW will be assessed. Because it would not be feasible to use the JA as well, the ACA was only compared with the DW. Since JA is rarely used and the scientific community has overwhelmingly embraced the DW, despite the lack of data on its validity, we wished to compare ACA and DW. Further, relationships of the resulting iQoL index scores with scores on a VAS for QoL and for iQoL may give more insight into the validity of both weighting methods.

Methods

Patients

To assess the feasibility and validity of the ACA to derive iQoL weights, a convenience sample of outpatients with rheumatoid arthritis or cancer who were treated at the Leiden University Medical Center were asked to participate in the study. We selected patients with rheumatoid arthritis (RA) who received multidisciplinary day treatment or had an appointment with the specialized nurse consultant about their treatment. Patients with cancer were selected if they received curative radiotherapy at the time of the study or had received curative radiotherapy in the 6 months before. The latter patients received a letter at home in which the head of the Department of Radiotherapy asked them to participate in the study. These groups were selected because they presented at the clinic with symptoms (RA) or were known to have side effects (cancer) impacting their quality of life. All patients were only included in the study after they had given their informed consent. The Medical Ethical Committee of the Leiden University Medical Center approved the research protocol.

Interview and questionnaire

Patients were interviewed either at the hospital or at home. First, patients had to rate their current global QoL on a horizontal VAS anchored at the two extremes by the terms ‘best imaginable quality of life’ and ‘worst imaginable quality of life’. Next, iQoL was assessed by the SEIQoL. If patients were unable to nominate five areas of life, they were presented with a list with predetermined domains, such as family, health, finances, living conditions, work, social, and leisure activities, to help them make a choice [3]. Next, patients rated their functioning on each of these domains on five adjacent 0–100 mm vertical VAS scales anchored at the two extremes by the terms ‘best possible’ and ‘worst possible’. Further, patients had to rate their global iQoL, given their ratings on the five domains, on a VAS anchored at the two extremes by the terms ‘the best life I can imagine’ and ‘the worst life I can imagine’. In this article, the score for global iQoL on the VAS is named as SEIQoL-VAS. Finally, the weights of the iQoL domains were determined from the DW and the ACA.

The DW-pie consists of five stacked, centrally mounted, interlocking laminated discs. Each disc has a different color and is labeled with one of the five domains nominated by the individual. The discs can be rotated over each other to produce a dynamic pie chart where the relative size of each sector represents the weight the respondent attaches to a QoL domain [1].

The ACA is a computerized questionnaire [18]. At the start of the ACA survey, respondents were asked to indicate, using the mouse or key strokes, how important they considered the difference between the best and the worst level of functioning on each of the five SEIQoL domains on a seven-point Likert scale, with adjectives at the first (not important), third (somewhat important), fifth (quite important), and seventh (very important) radio button. Next they were presented with a series of paired comparisons. The pairs consisted of two scenarios, one on the left-hand side, and one on the right-hand side (see appendix). The scenarios differed with respect to the level of functioning on two or three domains of QoL. Although the use of scenarios with four or five domains would have provided for more precise estimates, we had to balance this against feasibility. We decided such scenarios were too difficult for this first experience using the ACA to assess SEIQoL weights. Each domain had four levels of functioning: very well, fairly reasonable, rather bad, and very bad. Respondents had to indicate their preference for the alternative on the left or right on a nine-point Likert scale. The number of pairs presented was based on the formula $3 \times (N - n - 1) - N$, where N is the number of levels across all domains and n is the number of domains, resulting in $3 \times (20 - 5 - 1) - 20 = 22$ pairs [19]. This formula presents a rule of thumb leading to three times the number of observations as parameters available (for simulations on accuracy of prediction with various numbers of pairs, see [20]). We presented 25 paired comparisons, where scenarios were defined using two (pairs 1–15) or three (pairs 16–25) domains.

Finally, patients filled out a questionnaire that addressed demographic factors such as age, sex, marital status, education, and religion.

Computation of iQoL weights and index scores

For the DW, the relative weight of a domain is equal to the proportion of the pie chart that its sector represents, which can be read from a 100-point scale on the circumference. Relative weights for the ACA are calculated as follows. First, patient utilities for all levels of functioning on the domains are derived by ordinary least-squares regression analysis, from participants’ answers to the pairwise comparisons, assuming a linear main effects additive model (for details see [18]). Next, the relative weights for the domains are calculated by dividing the range of each domain (utility of highest level – utility of lowest level) by the sum of ranges of all domains, and multiplying by 100 [16, 19]. The relative weights are expressed as percentages (the five weights add up to 100%) and reflect the extent to which the difference between the best and worst levels of each domain drives the decision to choose a specific scenario [16].

The index score for iQoL is a weighted score, calculated by multiplying the functioning scores for the domains with their corresponding weights as derived by the DW and the ACA method, and summing these. Further, an unweighted index score was calculated by simply summing up the functioning scores and dividing by 5.

Feasibility and validity

The feasibility of the ACA was assessed by measuring the percentage of patients that were able to finish the task, by measuring the administration time, and by asking the patients how they evaluated the ACA with respect to difficulty and acceptability. We asked patients two quantitative items about the method being confronting (very, somewhat, not) or being unpleasant versus fun (1 = very unpleasant, 5 = much fun). Further, we also coded qualitative statements about the ACA being upsetting (comments such as ‘nasty’, ‘mean’, ‘suicide questions’, ‘I felt like a prisoner’). As a measure of difficulty we also assessed how often patients chose the worst option in a dominant pair, a pair in which one of the scenarios was on all domains better than the other.

The validity of the ACA was first studied by assessing the number of inconsistencies in the rank ordering of utilities, that is, the number of pairs in which the utilities for two levels of functioning were ranked opposite to the direction of the levels of functioning. We analyzed whether age, health status, and level of education were related to answers to dominant pairs and the number of inconsistencies by Pearson’s correlation and analysis of variance. Next, we assessed whether patients were willing to trade off a decrease from the best to the second-best functioning level on their most important domain with the largest improvement on their second important domain. This was done by computing the ratio between the difference in utilities for the largest benefit in the second important domain and the difference in utilities for the two highest functioning levels of the most important domain. A value smaller than 1 was taken as indicating that the patient was not willing to trade off decline in the most important domain for any benefit in the second important domain. We similarly assessed whether patients were willing to trade off a decrease from the second to the third functioning level on their most important domain with the highest improvement on their second important domain.

Agreement

The absolute differences between the weights assigned in the DW and in the ACA were described by means and standard deviations. Agreement between the weights was assessed by the intraclass correlation coefficient and

Pearson’s correlation coefficient. Pearson’s correlation represents the linear association between two measures and is not strictly a measure of agreement. However, a Pearson’s correlation which is much larger than the intraclass correlation would provide some indication of a systematic change between measures, as might occur if there were learning effects [21]. Correlations between the weighted scores and the unweighted score on the one hand, and the VAS scores for QoL and iQoL on the other hand, were assessed using Pearson’s correlation coefficient. We used SPSS version 12.0 for Windows.

Results

Patients

Of the 71 patients approached, 27 patients with cancer and 20 patients with rheumatoid arthritis were willing to participate (response rate 66%). Twenty cancer patients and four patients with rheumatoid arthritis declined to participate. Reasons not to participate were: too burdensome ($n = 5$), patients too busy with treatment or work ($n = 3$), and other ($n = 3$). In 13 cases the reason was unknown: two patients did not give a reason and 11 patients did not respond at all. Characteristics of the patients are described in Table 1. Patients were, on average, 61 years old (SD 10 years). In 32 cases (68%), the interview was held at the hospital, the other interviews were held at home.

Table 1 Characteristics of patients ($N = 47$)

	<i>N</i> (%)
Sex	
Female	24 (51)
Living arrangement	
With partner	41 (87)
Education	
Low ^a	19 (40)
Religion	
Religious	25 (53)
Diagnosis	
Rheumatoid arthritis	20 (43)
Breast cancer	11 (23)
Prostate cancer	11 (23)
Rectal cancer	5 (11)
Place of interview	
Hospital	32 (68)
At home	15 (32)

^a Low education: lower vocational, lower secondary general education, or primary school; high education: intermediate vocational, higher secondary general education, higher vocational education, or university

Table 2 Nominated cues during the first stage of the SEIQoL ($N = 47$ patients)¹

	<i>N</i> (%)
Partner	22 (47)
Children	12 (26)
Partner and children	13 (28)
Family	18 (38)
Own health	30 (64)
Health of partner	5 (11)
Social contacts and friendship	21 (45)
Transportation	10 (21)
Independence	5 (11)
Hobbies and relaxation	23 (49)
Work	21 (45)
Feelings	5 (11)
Activities of daily life	5 (11)
Sports and holidays	11 (23)
Other	34 (72)
Total	235

¹ Only domains that were mentioned by five or more patients are reported, the others (such as sexuality and income) are grouped together in the final row

ACA: feasibility and validity

All patients nominated five areas which they considered most important to their QoL, of whom two patients had to consult the list with predetermined domains. Domains of life mentioned most frequently were own health, relationships with partner and children, social contacts and friendships, hobbies and recreation, and work (Table 2).

The ACA survey took on average 20 min (range 10–37 min). All patients were able to finish the ACA. Five patients (11%) were in some sense upset about the ACA survey, and a further three (6%) judged the questions as very confronting. An additional patient found it very unpleasant. For example, when a patient had nominated

own health and relationship with the partner as domains, the ACA could offer one scenario in which the patient's health was very good whereas the relationship with the partner was poor, and another scenario in which the patient's health was very poor whereas the relationship with the partner was very good. Some patients became upset when they had to make a choice between such options.

Patients were offered, on average, 2.9 dominant pairs (range 0–6). In such pairs, four times (3%) the worst option was chosen and once (1%) the patient had no preference. The four patients who chose the worst option had the same level of education and were of the same age as the other patients.

When the utility assigned to a higher level of functioning on a particular domain is lower than that assigned to a lower level of functioning on that same domain, this is inconsistent. Because each domain had four functioning levels, six different pairs of levels $[(4 \times 3)/2]$ can be constructed for each domain, that is a total of 30 pairs of utilities per person. On average, patients' utilities were rank-ordered opposite to the level of functioning in 3.9 out of these 30 pairs (13%; Table 3). For the most important domain, the mean number of inconsistencies was 0.2, whereas this was 1.7 for the least important domain, and the correlation between the inconsistencies and the domain weights was $r = -0.50$ ($P = 0.000$, $n = 235$, 47×5 weights). The number of inconsistencies was lower in patients with a higher education level (Spearman's rho -0.30 ; $P = 0.04$), whereas it was not related to age or health status.

One out of 43 patients was not willing to trade off a decline from the best to second-best functioning level on the most important domain for the largest benefit on the second important domain. Further, all patients were willing to trade off a decline from the second to the third level of the most important domain for the largest benefit on the second important domain.

Table 3 Consistency in utilities for levels of functioning of individual quality of life domains^a

	Most important				Least important	Total $N = 47$
	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	
<i>Number of inconsistencies in rank order utilities out of six pairs per domain</i>						
Mean (SD)	0.2 (0.6)	0.2 (0.5)	0.8 (1.0)	1.0 (1.2)	1.7 (1.1)	3.9 (2.1)
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)
No inconsistencies	41 (87)	41 (87)	25 (53)	23 (49)	8 (17)	138 (59)
One inconsistency	2 (4)	4 (9)	7 (15)	9 (19)	11 (23)	33 (14)
Two inconsistencies	4 (9)	2 (4)	13 (28)	10 (21)	17 (36)	46 (20)
Three inconsistencies			2 (4)	4 (9)	9 (19)	15 (6)
Four inconsistencies				1 (2)	2 (4)	3 (1)

^a The content of the domains may vary between patients, according to what an individual patient evaluates as most important

Table 4 Absolute differences between DW weights and ACA weights

	Most important ^a				Least important ^a	Total (<i>N</i> = 235)
	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	
Mean (SD)	7.5 (5.3)	4.7 (4.2)	4.4 (3.5)	4.6 (4.1)	5.9 (4.9)	<i>P</i> = 0.003
Absolute difference between ACA and DW	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)
Less than 5 points	17 (36)	31 (66)	28 (60)	31 (66)	26 (55)	133 (57)
5–10 points	13 (28)	12 (26)	16 (34)	13 (28)	12 (26)	66 (28)
More than 10 points	17 (36)	4 (8)	3 (6)	3 (6)	9 (19)	36 (15)

^a Importance based on adaptive conjoint analysis procedure

Table 5 Agreement between DW weights and ACA weights

	ACA–DW linear correlation		ACA–DW intraclass agreement	
	Pearson <i>r</i>	<i>P</i>	ICC ^b	<i>P</i>
<i>Domain of individual quality of life</i> ^a				
Most important domain	0.27	0.06	0.23	0.06
Domain 2	0.30	0.04	0.28	0.03
Domain 3	0.43	0.003	0.33	0.01
Domain 4	0.33	0.02	0.28	0.03
Least important domain	0.22	0.14	0.18	0.11

ACA, adaptive conjoint analysis; DW, direct weighting

^a Rank ordered according to weighting derived by the method of adaptive conjoint analysis

^b Two-way mixed effects model where people effects are random and measure effects are fixed; *f*-test with true value 0

Agreement between DW weights and ACA weights

The mean absolute difference between the DW weights and the ACA weights varied from 4.4 to 7.5 for the five domains of iQoL, on a scale from 0–100 (Table 4). For all five domains together, 36 pairs (15%) differed by more than 10 points. For the most important domain according to the ACA weighting, 17 patients (36%) had a difference of more than 10 points between their DW weight and ACA weight.

The correlation between the DW weights and the ACA weights varied from 0.22 to 0.43, and the intraclass correlation coefficient varied from 0.18 to 0.33, both indicating a low agreement between the two weighting methods (Table 5).

Consequences of weighting method

The index score for iQoL calculated with the weights derived by the ACA (SEIQoL-ACA) was slightly higher than the index score of the SEIQoL-DW [mean score 71.6 (SD 11.5) versus 70.1 (SD 12.2); *P* = 0.02] and both were higher than the unweighted index (mean 67.2, SD 12.4). The SEIQoL-VAS score correlated strongly and almost equally with the SEIQoL-ACA, the SEIQoL-DW, and the unweighted score. The three index scores were also positively related to the global QoL VAS (Table 6).

Table 6 Impact of weighting procedure on index score for individual quality of life, and on correlations with global quality of life

	DW index <i>r</i>	ACA index <i>r</i>	Unweighted index <i>r</i>
<i>Index score for individual quality of life</i>			
DW-index score	1.0	0.95**	0.92**
ACA-index score	0.95**	1.0	0.89**
Unweighted index score	0.92**	0.89**	1.0
SEIQoL VAS	0.62**	0.54**	0.63**
QoL VAS	0.40*	0.36*	0.33*

P* < 0.05; *P* < 0.001

DW, direct weighting; ACA, adaptive conjoint analysis; SEIQoL, schedule for evaluation of individual quality of life; VAS, visual analogue scale; QoL, quality of life

Discussion

We explored whether the ACA was a feasible and valid method to derive weights for iQoL domains. The iQoL community seems to have fully embraced the DW, despite its lack of agreement with the original JA method. Contrary to JA, the DW has no theoretical underpinning. Further, DW is a direct method, and as such may not capture implicit or unconscious thoughts or preferences. Since respondents will generally not be explicitly aware of the contributions of their life domains to their overall QoL,

indirect methods may be of more value. JA is not feasible, however, in most situations, and for this reason we piloted another theory-based indirect method, the ACA. ACA and JA are based on the same premises. However, JA was initially developed for assessing expert opinion [22, 23], respondents have to rate many scenarios, and the task is cumbersome. Due to its adaptive nature, ACA is much less cumbersome than a full-scale conjoint analysis, which is a nonparametric form of JA. In addition to being indirect, ACA can provide measures of internal reliability and validity (inconsistencies, willingness to trade) for individual interviews. The agreement between DW weights and ACA weights may give insight into the validity of the two methods.

ACA was to some extent feasible, because the ACA took on average 20 min and all patients were able to finish it. However, one in five patients judged the ACA task as upsetting, very confronting, or very unpleasant. The paired comparison task, despite working well for some domains, turned out not to be appropriate for some others. Especially choosing between two domains that are dear to the patient turned out not to be feasible. Sometimes, the computer offered a dominant pair, mostly resulting from the fact that the utilities of the functioning levels were almost equal. Only seldom was the worst option chosen. This finding shows that almost all patients understood the task of paired comparisons and were able to make a valid choice.

A limitation of our procedure was that patients did not rate scenarios with four or five domains. Although the use of such scenarios would have provided for more precise estimates, this had to be balanced against feasibility. Using all five attributes in the pairwise comparisons would also have allowed for the evaluation of full profiles (health states). The goal of this first study on the use of ACA for the SEIQoL was merely to assess its feasibility and to compare ACA weights with DW weights, not to assess full profiles. We therefore preferred to opt for the more feasible approach, which we deemed sufficiently difficult already.

Many patients gave inconsistent answers, but these inconsistencies mostly occurred on domains 3 and lower, and especially on the least important domain. For the two most important domains, the large majority of patients had utilities ordered in the same direction as the corresponding levels of functioning. In the case of less important domains, the differences between utilities of successive functioning levels are probably small, which leads to inconsistent answers.

A major premise of the ACA is that respondents are willing to use compensatory decision making. Our findings show that almost all people were willing to do so. The willingness to trade off the difference between a decline on the most important domain with the largest benefit on the second most important domain is consistent with the

finding that, in patients with colorectal cancer, 95% were willing to trade off a 1% loss of survival [17].

The agreement between DW and ACA weights was low. Browne et al. reported similar findings for the comparison between JA and DW [1]. Further, the weights derived from each method were only poorly or moderately correlated to each other. The differing weights from the DW and the ACA suggest that the two methods are not interchangeable.

Weighting had almost no effect on relationships with other global measures of QoL, and the DW index score correlated 0.95 with the ACA index score, and both weighted index scores were also highly correlated to the unweighted index score. Weighting or not weighting domains of QoL has received much attention in the literature. Some studies have shown that there is no effect of weighting [24–28]. For iQoL, the findings of Wettergren et al. suggest that the degree of importance is already built into the process of selection, when the participants themselves select the domains [28]. Further, it has been suggested that a satisfaction evaluation had incorporated the judgment of item importance [29]. Our findings indeed show that the weighted iQoL index scores were higher than the unweighted index score, which indicates that patients gave higher weights to domains with better functioning. Most likely, the patients already take into account the importance of a specific domain, first in the selection of the domain and next when they evaluate the functioning on that domain.

Our study has some limitations. The number of patients was relatively small, due to the qualitative character and semistructured design of the SEIQoL interview. However, larger numbers would not have changed the conclusion of our paper. A problem in measuring the validity of the SEIQoL is that a gold standard for iQoL is lacking. JA has been considered the standard weighting method, but due to its complex nature has been replaced in the field by DW. Unfortunately, it was not possible to include both the JA and the ACA, because of the cognitive burden imposed on the patients.

Despite the patients' reluctance to perform the ACA, our study gave clear insight into the problems of deriving weights for iQoL domains. Our findings show that weighting has almost no effect on the association between the SEIQoL and global iQoL, although incorporating weights for domain functioning led to slightly higher iQoL index scores than the unweighted index score. Selecting and weighting domains are clearly confounded. Because of the high correlations between the weighted and unweighted index scores, it seems sufficient to use the unweighted index score as a measure for global iQoL.

Acknowledgements A.M. Stiggelbout and E. de Vogel-Voogt were supported entirely by a *VIDI-award* of the Netherlands Organization

for Scientific Research NWO Innovational Research Incentives (grant number 917.56.356). The authors would like to thank the patients who participated in this study and especially R.M. Westerman for her help in selecting cancer patients who were eligible for the study. Further, we gratefully appreciate the assistance of A.E.M. Jongma, E.G.M. Kolman, the health care professionals of the day-care center Sole Mio, and the radiotherapists who invited patients to participate in the study. We thank Nanny van Duijn for her help with data collection.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix

Example of a pairwise comparison screen in ACA

All other domains being equal, which of the two situations below would you prefer?

<p>You live in the following situation:</p> <ul style="list-style-type: none"> •Regarding family things go very well •Regarding your own health things go rather bad 	<p>You live in the following situation:</p> <ul style="list-style-type: none"> •Regarding family things go fairly reasonable •Regarding your own health things go fairly reasonable
--	---

0 0 0 0 0 0 0 0 0 0

strong preference for left hand
strong preference for right hand

References

- Browne, J. P., O'Boyle, C. A., McGee, H. M., McDonald, N. J., & Joyce, C. R. (1997). Development of a direct weighting procedure for quality of life domains. *Quality of Life Research*, 6(4), 301–309.
- McGee, H. M., O'Boyle, C. A., Hickey, A., Omalley, K., & Joyce, C. R. B. (1991). Assessing the quality-of-life of the individual - the SEIQoL with a healthy and a gastroenterology unit population. *Psychological Medicine*, 21(3), 749–759.
- O'Boyle, C. A., McGee, H., Hickey, A., O'Malley, K., & Joyce, C. R. (1992). Individual quality of life in patients undergoing hip replacement. *Lancet*, 339(8801), 1088–1091.
- Clarke, S., Hickey, A., O'Boyle, C., & Hardiman, O. (2001). Assessing individual quality of life in amyotrophic lateral sclerosis. *Quality of Life Research*, 10(2), 149–158.
- Echteld, M. A., Deliens, L., Ooms, M. E., Ribbe, M. W., & van der Wal, G. (2005). Quality of life change and response shift in patients admitted to palliative care units: A pilot study. *Palliative Medicine*, 19(5):381–388.
- Lindblad, A. K., Ring, L., Glimelius, B., & Hansson, M. G. (2002). Focus on the individual—quality of life assessments in oncology. *Acta Oncologica*, 41(6), 507–516.
- Wettergren, L., Bjorkholm, M., Axdorph, U., Bowling, A., & Langius-Eklöf, A. (2003). Individual quality of life in long-term survivors of Hodgkin's lymphoma—a comparative study. *Quality of Life Research*, 12(5), 545–554.
- McGee, H., Hickey, A., Oboyle, C., Omalley, K., & Joyce, C. R. B. (1988). Judgment analysis (ja) as a technique for providing individualized assessments of quality of life (QoL). *British Journal of Clinical Pharmacology*, 26(5), 653–654.
- Joyce, C. R., Hickey, A., McGee, H. M., & O'Boyle, C. A. (2003). A theory-based method for the evaluation of individual quality of life: The SEIQoL. *Quality of Life Research*, 12(3), 275–280.
- Waldron, D., O'Boyle, C. A., Kearney, M., Moriarty, M., & Carney, D. (1999). Quality-of-life measurement in advanced cancer: Assessing the individual. *Journal of Clinical Oncology*, 17(11), 3603–3611.
- Schwarz, N. (1999). Self-reports—how the questions shape the answers. *American Psychologist*, 54(2), 93–105.
- Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, 8(3), 355–363.
- Louviere, J. J. HDASJD, & Adamowicz, W. (2000). *Stated choice methods: Analysis and application*. Cambridge: Cambridge University Press.
- Ryan, M., & Farrar, S. (2000). Using conjoint analysis to elicit preferences for health care. *BMJ*, 320(7248), 1530–1533.
- Beusterien, K. M., Dziekan, K., Flood, E., Harding, G., & Jordan, J. C. (2005). Understanding patient preferences for HIV medications using adaptive conjoint analysis: Feasibility assessment. *Value in Health*, 8(4), 453–461.
- Fraenkel, L., Bodardus, S., & Wittnik, D. R. (2001). Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. *Medical Care*, 39(11), 1203–1216.
- Pieterse, A. H., Stiggelbout, A. M., Baas-Thijssen, M. C. M., Van de Velde, C. J. H., & Marijnen, C. A. M. (2007). Benefit from pre-operative radiotherapy in rectal cancer treatment: Disease-free patients' and oncologists' preferences. *British Journal of Cancer*, 97(6), 717–724.
- ACA 5.0 Technical paper (2005). Sawtooth Software Technical paper series. 13-9-2007. Ref Type: Electronic Citation.
- Orme, B. (2005). *SSI Web v5. Software for web interviewing and conjoint analysis*. Sequim, WA: Sawtooth Software, Inc.
- Johnson, R. (2007). Accuracy of utility estimation in ACA. Technical paper. Sawtooth Software Technical paper series. 14-9-2007. Ref Type: Electronic Citation.
- Shiell, A., Seymour, J., Hawe, P., & Cameron, S. (2000). Are preferences over health states complete? *Health Economics*, 9(1), 47–55.
- Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 271–307). New York: Academic Press.
- Stewart, T. R., & Joyce, C. R. B. (1988). Increasing the power of clinical trials through judgment analysis. *Medical Decision Making*, 8(1), 33–38.
- Gorbatenko-Roth, K. G., Levin, I. P., Altmaier, E. M., & Doebbeling, B. N. (2001). Accuracy of health-related quality of life assessment: What is the benefit of incorporating patients' preferences for domain functioning? *Health Psychology*, 20(2), 136–140.
- Jansen, S. J., Stiggelbout, A. M., Nooij, M. A., & Kievit, J. (2000). The effect of individually assessed preference weights on the relationship between holistic utilities and nonpreference-based assessment. *Quality of Life Research*, 9(5), 541–557.

26. Trauer, T., & Mackinnon, A. (2001). Why are we weighting? The role of importance ratings in quality of life measurement. *Quality of Life Research*, 10(7), 579–585.
27. Wu, C. H., & Yao, G. (2006). Do we need to weight satisfaction scores with importance ratings in measuring quality of life? *Social Indicators Research*, 78(2), 305–326.
28. Wettergren, L., Bjorkholm, M., & Langius-Eklöf, A. (2005). Validation of an extended version of the SEIQoL-DW in a Cohort of Hodgkin Lymphoma' survivors. *Quality of Life Research*, 14(10), 2329–2333.
29. Wu, C. H., & Yao, G. (2007). Importance has been considered in satisfaction evaluation: An experimental examination of Locke's range-of-affect hypothesis. *Social Indicators Research*, 81(3), 521–541.