doi: 10.1111/j.1369-7625.2010.00595.x

Methodologic evaluation of adaptive conjoint analysis to assess patient preferences: an application in oncology

Arwen H. Pieterse PhD,* Anne M. Stiggelbout PhD† and Corrie A.M. Marijnen PhD‡

*Dutch Cancer Society postdoctoral fellow, †Professor of Medical Decision Making, Department of Medical Decision Making, Leiden University Medical Center, Leiden, The Netherlands and ‡Professor of Radiotherapy and Head of the Department of Radiotherapy, Department of Clinical Oncology, Leiden University Medical Center, Leiden, the Netherlands

Correspondence
Arwen H. Pieterse PhD
Department of Medical Decision Making
Leiden University Medical Center
PO Box 9600
2300 RC Leiden
The Netherlands
E-mail: a.h.pieterse@lumc.nl

Accepted for publication
7 December 2009

Keywords: computer-assisted interviewing, conjoint analysis, rectal cancer, utility assessment

Abstract

Background Adaptive conjoint analysis (ACA) is an individually tailored preferences elicitation technique that mimics actual decision-making processes by asking participants to make trade-offs between the various dimensions that underlie decision problems. ACA is increasingly applied in patient preferences assessments but formal evaluation of its validity and reliability is lacking.

Objective To investigate ACA's validity and reliability in elicitation of treatment outcome preferences.

Methods Sixty-eight disease-free rectal cancer patients, treated with surgery with or without preoperative radiotherapy were asked to complete exercises to assess their preferences for radiotherapy [using the treatment trade-off method (TTM)] and for key outcomes associated with radiotherapy (using ACA). We assessed (i) rank ordering of ACA-derived outcome-probability utilities, (ii) compensatory decision making, (iii) ACA test–retest reliability, and (iv) concordance of ACA- and TTM-based preferences.

Results All participants completed the TTM and 66 completed the ACA questionnaire, in 15 min on average. Outcome utilities were rank ordered in agreement with probabilities from best to worst in most participants, except for sexual dysfunction. Most participants were willing to trade survival and their most important outcome. Mean importance ratings were similar at retest. ACA- and TTM-based preferences differed. TTM-based preferences were related to past treatment, ACA-based preferences were not.

Conclusions ACA assesses group-level preferences reliably over time and captures individual preferences independently from treatment experience in treated cancer patients. ACA seems a valid treatment outcome preference elicitation method in a context in which trade-offs between cure and quality of life need to be considered.

 © 2010 The Authors. Health Expectations © 2010 Blackwell Publishing Ltd

Introduction

For many health problems, no single treatment choice dominates because, for example, treatment carries significant risks or the decrease in risk of negative outcomes associated with an alternative treatment is marginal.[1,2] In these situations, clinicians cannot justifiably tell patients which is the 'best' treatment. Also, physicians show inaccuracies at judging individual patients' values for particular outcomes of care,[3–5] and at predicting their treatment preferences.[6] It is therefore essential that individual patients participate in choosing the treatment that best fits them. To do this, they must make trade-offs between the relevant dimensions of treatment options.

Involving patients in treatment decision making has been shown to improve decision quality.[7] However, patients are often unfamiliar with the options and outcomes considered in treatment decisions, so they will lack ready-made insight into their valuation of relevant dimensions.[8] Value assessment techniques may help them to acquire such insight and may thus aid patient decision making as well as inform the formulation of treatment guidelines.

A promising technique to clarify personal values patients associate with treatment dimensions is adaptive conjoint analysis (ACA), which has been developed as a more feasible alternative to traditional conjoint analysis. Like traditional conjoint tasks, ACA is a decomposed, and therefore indirect, method, eliciting preferences regarding distinct dimensions associated with treatment options. The exercise primarily asks participants to rate combinations of treatment dimensions that involve trade-offs (typically because a gain in one dimension is related to loss in another; Fig. 1). The trade-offs are similar to the way individuals make decisions in real life. Conjoint analysis is said, therefore, to be more effective for identifying true preferences than holistic techniques.[9,10] The main difference between ACA and traditional CA is that the computerized administration allows for tailoring of the combinations offered to the individual's preferences. As the task progresses, combinations of dimensions become more and more equal in desirability for the participant. ACA thus facilitates participants' involvement[11] and also drastically reduces the number of combinations that participants need to evaluate.[11] Of note, tailoring inherently makes the trade-offs increasingly difficult. The standardized computer administration minimizes influences on participant preferences.

ACA has other advantages over conventional preference assessment techniques. Unlike techniques such as Standard Gambles and Time TradeOffs, ACA permits the evaluation of dimensions simultaneously and of both chronic

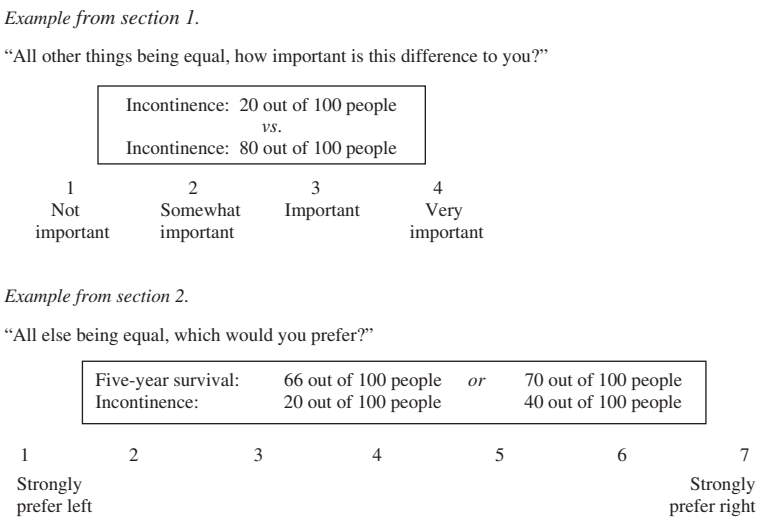

Figure 1 Adaptive conjoint analysis questionnaire.

and temporary states. Differently from decision boards,[12,13] videotapes[14,15] and rating scales, which usually ask about preferences for treatments rather than valuations of underlying dimensions, ACA provides insight into the basis of preferences. Results therefore extend beyond particular medical conditions.

The ACA has a firm theoretical basis;[16] it allows for the assessment of response consistency and thus permits a methodological evaluation of results; and it supports decision-making processes in a way that shows to full advantage actual decision problems, by presenting the trade-offs underlying them. Accordingly, ACA is increasingly being used to investigate group-level treatment preferences, as is demonstrated by studies in rheumatology patients,[9,17,18] HIV patients,[19] patients using hormone growth therapy[20] and patients about to undergo major abdominal surgery.[10] Mimicking actual decision making processes, being relatively short and individually tailored, ACA could shed light on individual patients' health-related priorities and thus support their decision making.

Currently there is little information available on ACA's methodologic properties in the context of health care, except for two studies. One study looked at the external validity of ACA-derived utilities by assessing their consistency with preferences for medication profiles in HIV patients.[21] The other study compared weights for individualized quality of life domains derived from direct weighting vs. ACA, in rheumatology and cancer patients.[22] The latter study additionally assessed ACA completion, response consistency, and compensatory decision making (i.e. it examined participants' willingness to trade loss in one dimension for benefit in another: the assumption of compensatory decision making would be violated if individuals determined their preference by considering outcome dimensions in order of importance).[23]

To our knowledge, no study has assessed ACA test–retest reliability, or compared ACA to alternative preference assessment methods. The present study was designed as a first step in the methodological assessment of ACA as a treatment outcome preference elicitation method in cancer patients. The ACA work could support two ultimate goals: to support newly diagnosed patients in practice and to inform treatment guideline formulation. As clinical case, we chose to assess disease-free rectal cancer patients' preferences for outcomes related to two options: surgery vs. preoperative radiotherapy (PRT) followed by surgery for rectal cancer. These two treatments have distinct risk/benefit ratios; some side effects are enduring and may have a profound influence on quality of life.

We examined (i) the rank order consistency of ACA-derived outcome-probability utilities, as an indication of response consistency; (ii) participants' willingness to trade-off outcome dimensions, as an indication of compensatory decision making and thereby of internal validity; (iii) ACA test–retest reliability, as an indication of response stability over time; and (iv) concordance of ACA- vs. treatment trade-off method (TTM)-based treatment preferences, as an indication of external validity. The TTM is a direct and holistic preference elicitation technique that is tailored to the clinical problem at hand. The TTM has been applied extensively to elicit patient treatment preferences, including preferences in clinical oncology practice.[12,13,24,25] It was shown to have fair to good test–retest reliability.[26,27]

Differently from other ACA applications in health care, we included survival as a dimension. Survival has been incorporated in studies using traditional conjoint methods,[28–30] but not yet in ACA. It is unclear to what extent patients, particularly cancer patients, are willing to trade-off survival when the exercise is increasingly customized to individuals' priorities.

Methods

Participants

Participants were disease-free rectal cancer patients who had participated in a multicentre trial assessing the benefit of adding PRT to surgery.[31] A stratified random sample was selected from the patients who had been treated in the Netherlands, had agreed to being

approached for further research at 5-year follow-up,[32] were alive and free of recurrences or other tumours at last moment of yearly follow-up, were less than 90 years old when our inclusion started, and had undergone low anterior (LAR) or abdominoperineal (APR) resection. Sampling was stratified to obtain as much variance in preferences and trade-offs as possible. We aimed to include 60 participants, half of whom had been treated with surgery alone and half also with PRT. We further aimed to include both patients who did and who did not face sexual problems and/or faecal incontinence, two longer-term quality of life impairments which may follow rectal cancer treatment. We took equal numbers of patients who had vs. had not reported sexual dysfunction at 2-year follow-up[33] in both treatment groups. We further stratified the treatment groups according to surgery type: two-third had undergone an LAR and one-third an APR, the latter resulting in a permanent stoma. LAR patients face the risk of faecal incontinence. We took equal numbers of LAR patients who had vs. had not reported faecal incontinence at 5-year follow-up.[32] Eligible participants were informed about the study by letter and then asked by phone whether they agreed to participate. Participants were included between February and August 2006.

Trial results

Six-year follow-up data showed PRT followed by surgery to result in a reduction in the local recurrence rate from 11% to 6% with no survival benefit.[34] Two- and 5-year follow-up results showed both treatments to induce faecal incontinence and sexual dysfunction, yet probabilities were higher with PRT.[32,33]

Procedure

Socio-demographic and medical history details were assessed by self-report questionnaire a week prior to a face-to-face interview at participants' home. Patients who did not have a stoma were asked how often (never, sometimes, often, always) they had unintentionally lost stools in the past 4 weeks. Male (erectile and ejaculation dysfunction) and female (vaginal dryness and pain) sexual problems in the past 4 weeks (not at all, a little, quite, a lot) were assessed as part of the EORTC QLQ-CR38.[35] Female sexual dysfunction was only assessed in patients who had indicated to have had intercourse in that period.

At the start of the interview, participants were asked for informed consent. A TTM and ACA questionnaire were administered in that order. The ACA questionnaire was computerized and self-administered. We expected the ACA task, which asks for reflection about the underlying dimensions, to influence responses to the TTM task more than the other way round. Also, we did not want to interrupt the flow of the interaction by introducing a computer task midway. A convenience sample of $N = 30$ consecutive participants was asked to perform the ACA questionnaire again, 2–4 weeks later. Reasons for refusal were noted. The Leiden University Medical Center ethical board approved the study protocol.

Treatment trade-off method

The TTM presented sequentially (i) a short description of surgery alone and PRT followed by surgery, (ii) the respective probabilities of faecal incontinence[32] and sexual dysfunction[33] following treatment, and (iii) the probability of 5-year local control[34] after surgery alone (Table 1). Male and female TTM versions differed regarding probabilities of sexual dysfunction. Probabilities were described in quantitative frequency formats to facilitate understanding.[36] It was explicitly stated that survival was similar following both treatments. Then probabilities for 5-year local control with PRT were varied and participants were asked each time which treatment they preferred.

We started with equal probabilities of local control after the two treatments. To participants who preferred PRT followed by surgery at this stage, we explained the numbers again. We stopped at this point if participants maintained a preference for PRT. To the others, we asked their preferences for the following probabilities:

Table 1 Treatment outcome probabilities as presented in the TTM (frequencies out of 100 patients)

Outcome	Treatment	
	Surgery	PRT + surgery
Faecal incontinence	40	60
Sexual dysfunction		
Men	56	68
Women[1]	15	22
Five-year local control	89	–[2]

TTM, treatment trade-off method; PRT, preoperative radiotherapy. Probability of survival was not presented.
[1]Numbers refer to sexually active women and this was stated as such.
[2]Probability of local control for PRT followed by surgery was systematically varied between 89/100 and 100/100.

5-year local control in 100/100 patients after PRT (i.e. maximum benefit from PRT) and 5-year local control in 95/100 patients after PRT. Next, we sought participants' minimally desired benefit by systematically closing in upon participants' indifference point simultaneously[37] from 89/100 (no benefit from PRT) and 100/100.

Adaptive conjoint analysis

We used the ACA questionnaire to ascertain participants' utilities for various probabilities of 5-year survival, 5-year local control, faecal incontinence, and sexual dysfunction. It was explicitly stated that survival and local control should be viewed as independent outcomes. Male and female ACA versions differed regarding sexual dysfunction. Probabilities were described in quantitative frequency formats. Treatment modality was not specified.

Each outcome was defined by various levels that corresponded to a range of probabilities, including those that were established at 2-,[33] 5-,[32] and 6-year[34] follow-up (Table 2). Probabilities for sexual dysfunction were close but not identical to those in the TTM, in order (i) to have some levels identical for men and women, increasing their comparability, and (ii) to include the percentages of dysfunction for sexually active patients, and for the whole group.

Table 2 ACA outcomes and outcome probabilities (frequencies out of 100 patients)

Outcome	Explanation	Outcome probabilities (from best to worst)			
Probability of 5-year survival (all patients)	This is the probability that the patient is still alive 5 years after the disease was detected. A 5-year survival of 50% means that after 5 years, 50 out of 100 patients are still alive. The other 50 people may have died due to the recurrence of the disease, but may as well have died from other causes such as a heart attack	70	66	65	–
Probability of 5-year local control[1] (all patients)	This is the probability that the tumour does not recur at the site that was operated on. If the tumour does recur at that site, it causes a lot of pain. It may in some instances be possible to treat it, but in others not. Often the prognosis is uncertain	99	94	89	–
Probability of faecal incontinence (all patients)	Incontinence in this interview refers to incontinence for stools and means unintentionally losing stools	20	40	60	80
Probability of sexual dysfunction (men)	You may think of dysfunction with getting an erection (=erectile dysfunction) and with ejaculation, or of not being sexually active at all anymore	30	40	50	60
Probability of dissatisfaction with sexual functioning (women)	Dissatisfaction with sexuality usually results from not being able to enjoy sexual intercourse anymore because of pain or vaginal dryness	10	30	50	70

ACA, adaptive conjoint analysis.
[1]The expression 'probability of local control' was not used but was explained as 'probability that the tumour does not recur'.

The ACA questionnaire started with an explanation in lay terminology of the treatment outcomes (Table 2) and asked participants to rate for each how important they considered the difference between its best and worst probability (all else being equal; Fig. 1) in the following order: survival, local control, sexual dysfunction and incontinence. This exercise provides the ACA programme with information upon which to base initial estimates of participants' utilities.[11]

Participants were then asked to rate 14 paired combinations of outcomes, where each combination was defined using two (pairs 1–7) or three (pairs 8–14) outcomes. The minimum number of pairs to be presented to permit accurate estimations of outcome-probability utilities, was based on the following formula: $3 \times (N - n - 1) - N$,[11] where N is the total number of probability levels, n the number of outcomes and $N - n$ the number of independent parameters to be estimated, thus resulting in $3 \times (14 - 4 - 1) - 14 = 13$ pairs. We started with the simplest paired comparison, i.e. pairs of two outcomes, offering participants the possibility to gain experience in completing the task. We chose to increase the number of outcomes per pair to three at most, in order to better mimic the actual trade-off – which implicates more than two outcomes – while limiting task difficulty. Participants were asked to indicate their preference on a 7-point scale (Fig. 1). Outcomes were listed in varying order. Usually, participants are also asked to rank levels for preference. We omitted this question because preferences over the outcome-probabilities are naturally ordered.

The ACA analysing program (ACA 5.2.2; Sawtooth Software, Inc., Sequim, WA, USA) generates estimates of individual participants' utilities for each outcome-probability. Final utilities are derived by ordinary least squares regression analysis from the participant's answers to the paired comparisons, and are continually re-estimated as the questionnaire progresses[11] (for detailed information on utility estimations, we refer the reader to: http://www.sawtoothsoftware.com/download/techpap/acatech.pdf). Similarly, the program computes participants' relative importance score for each outcome. Importance is expressed as a percentage, and the four importance scores add up to 100%. The scores indicate the extent to which one outcome explains participants' preference relative to the other outcomes, given the range of outcome-probabilities.

Statistical analysis

Scores for faecal incontinence were recoded to indicate no vs. at least some incontinence. Similarly, sexual dysfunction was coded as present in participants who had indicated to suffer at least 'a little' from one or both of the two problems assessed. We compared irradiated to non-irradiated patients on background (age, gender, education) and treatment-related (time since surgery, stoma, incontinence, sexual problems) characteristics using independent *t*-tests and chi-square tests.

As an indication of ACA feasibility, we examined the number of participants who completed the ACA questionnaire and time needed for completion. Task comprehension was evaluated by assessing preference in case of dominant paired comparisons.

As an indication of response consistency, we evaluated whether better outcome-probabilities generated higher estimated utilities than worse probabilities. To that end, higher probabilities of a good outcome or lower probabilities of a bad outcome will be referred to as 'better' probabilities. Lower probabilities of a good outcome or higher probabilities of a bad outcome will be referred to as 'worse' probabilities. The value 0 was assigned to the least preferred probability within each outcome. We computed how often estimated utilities were consistently rank ordered within each outcome. In participants who did not rate the worst probability highest, we assessed the association between rank ordering consistency and, respectively, age and outcome importance using Pearson r correlations, and education, using ANOVA and Kruskal–Wallis tests. Age and education were used as proxies for cognitive ability.

We examined whether participants' decision making was compensatory by inspecting participants' willingness to trade a loss in one outcome for improvement in another. Importantly, participants' willingness to trade will depend on the range of outcome-probabilities offered. This is especially relevant given the restricted range for survival and local control (Table 2). We chose to keep probabilities close to what would be clinically feasible, because of our ultimate goal to assess the usefulness of ACA in a clinical setting. We assessed participants' willingness to trade an outcome for the largest benefit in any of the other outcomes for the following two possibilities:

1. the smallest decrease in 5-year survival (i.e. the difference between the utilities for 66/100 and 65/100 patients alive);
2. a decrease from the best to the next-best probability in the outcome that the participant considered most important.

A ratio greater than 1 between the differences in utilities indicates that participants were not willing to trade a loss in one outcome for benefit in the other outcome. We chose to focus these analyses on the trading of survival and of participants' most important outcome, because these are the outcomes in which one would least expect compensatory decision making. Moreover, we conducted these analyses in participants whose outcome-probability utilities were consistently rank ordered in each respective pair of outcomes for reasons of interpretability. It is only in these participants that it is unequivocal which differences in outcome-probabilities are relevant for analysis.

We evaluated the test–retest reliability of ACA findings by comparing ACA test and retest data on importance scores, using paired *t*-tests. As ACA is adaptive, participants did not complete the exact same questionnaire at retest.

Lastly, we assessed the concordance of ACA- vs. TTM-derived treatment preferences. Participants' ACA-based treatment preferences were computed by adding their utilities for local control, incontinence and sexual dysfunction for the treatment-specific probabilities (Table 1). For sexual dysfunction, we used the utilities of the probabilities that were closest to the numbers shown in the TTM. Utility for survival was not included, because in explaining the TTM we stated that survival was shown not to differ between the two treatments. Participants' TTM-based treatment preference was inferred from their minimally desired benefit from PRT followed by surgery compared to surgery alone. If this benefit exceeded an absolute 5% increase in probability of local control, TTM-based preference was said to be surgery alone. This cut-off point was derived from the actual benefit from PRT that was established at 6-year follow-up. ACA- and TTM-based treatment preferences were compared using a McNemar test. For both it was assessed whether preference was dependent on having received PRT using a chi-square test. All significance testing was performed two-tailed at $\alpha = 0.05$.

Results

Participants

Ninety-four eligible patients were approached. Four could not be reached and nine had (had) other types of cancer or a recurrence and were therefore excluded. Of the remaining 81 patients, 70 (86%) agreed to participate. Eleven declined participation because they considered it too great a psychological burden ($N = 8$), physical burden ($N = 1$), time investment ($N = 1$) or for unknown reasons ($N = 1$). Table 3 shows demographic and treatment-related details by actual past treatment. In non-stoma patients, respectively 6/22 (27%) non-irradiated and 12/22 (55%) irradiated patients reported to suffer at least sometimes from faecal incontinence. The two participant groups did not significantly differ on background characteristics (data not shown).

ACA completion

Two male participants were not asked to complete the ACA questionnaire because of a computer failure (1) or having such difficulty in

Table 3 Patient demographic and treatment-related details by actual treatment received (N = 70)

Characteristics	Past treatment: Surgery, N (%)	Past treatment: PRT + surgery, N (%)
Mean age, years ± SD (range)	64.5 ± 9.8 (41–84)	63.8 ± 9.1 (43–82)
Gender		
Male	18 (56)	30 (79)
Female	14 (44)	8 (21)
Education		
≤9 years	12 (38)	19 (50)
10–12 years	15 (47)	9 (24)
≥13 years	5 (16)	10 (26)
Mean time since surgery, years ± SD (range)	8.4 ± 1.0 (6.3–10.0)	8.3 ± 1.0 (6.5–10.2)
Permanent stoma		
No	22 (69)	22 (58)
Yes	10 (31)	16 (42)
Sexual problems[1,2]		
No	2 (8)	9 (27)
Yes	22 (92)	24 (73)
Total	32 (46)	38 (54)

[1]In women, presence or absence of problems is only reported for the 6/14 (surgery) and 5/8 (PRT + surgery) who had been sexually active in the past 4 weeks.
[2]Data were missing for two men in the PRT + surgery group.

comprehending numbers that the TTM already was too challenging a task (1). Of the other 68 participants, one male and one female participant found the task too difficult and did not finish it. One stated: 'My head is spinning, I get fed up with it'. The other said to dislike the task, and that it required taking in a lot of information.

Completion took between 7 and 55 min (median = 14, SD = 7.3). The participant who took 55 min was dyslexic and reading the ACA instructions required extensive efforts. Completion took between 7 and 38 min if we disregard this participant.

Thirty-two of 66 participants were presented with at least one dominant pair, i.e. 17/32 once, 12/32 twice and 3/32 three times. They preferred the best option in 40/50 (80%), the worst in seven (14%) and indicated no preference in three (6%) cases. The seven worst option choices corresponded to seven different participants.

Table 4 Number of participants with consistent[1] rank orders for outcome-probability utilities (N = 66)

Outcome	Utilities in consistent rank order: N	Utilities in consistent rank order: %
Five-year survival	47	71
Five-year local control	58	88
Faecal incontinence	38	58
Sexual dysfunction	23	35
Men	19/45	42
Women	4/21	19

[1]Consistency was defined as adaptive conjoint analysis-derived utilities being rank ordered in agreement with probabilities from best to worst.

Rank ordering of estimated outcome-probability utilities

Table 4 shows how often outcome-probabilities were rank ordered in agreement with probabilities from best to worst, per outcome. In eight participants, utilities were rank ordered consistently in all four outcomes. In none, inconsistencies occurred in all four outcomes. Across outcomes and out of a total of 18 within-outcome pairs of outcome-probabilities, utilities were rank ordered inconsistently between 2.3 (13%) pairs on average (median = 2; SD = 1.6; range, 0–7). The highest utility estimation was conferred to the *best* probability in 61 (92%) participants for survival, 63 (95%) for local control, 54 (82%) for incontinence and 47 (71%) for sexual dysfunction. The highest utility estimation was conferred to the *worst* probability in zero participants for survival, one (2%) for local control, two (3%) for incontinence and one (2%) for sexual dysfunction. In one participant, all worse probabilities had higher estimated utilities than better probabilities for 'incontinence'. In another participant, this was true for 'sexual dysfunction'.

Excluding the participants who rated the worst probability highest, inspection of inconsistent pairs revealed that inconsistencies occurred most often between consecutive levels, i.e. pairs of probabilities with the smallest absolute differences (Appendix A). Higher consistency was significantly related to lower

Table 5 Association between frequency of outcome-probability utilities rank order consistency, and age and outcome importance (N = 66)

	Age		Importance	
Outcome	Pearson r	P-value	Pearson r	P-value
Five-year survival	0.06	0.66	0.32	0.01
Five-year local control[1]	−0.29	0.02	0.14	0.26
Faecal incontinence[2]	0.03	0.80	0.48	0.00
Sexual dysfunction				
Overall[3]	0.16	0.20	0.56	0.00
Men (N = 44)[3]	0.13	0.41	0.62	0.00
Women (N = 21)	0.22	0.34	0.31	0.17

[1]One inconsistent participant was excluded because he conferred the highest utility to the worst probability.
[2]Two inconsistent participants were excluded because they conferred the highest utility to the worst probability.
[3]One inconsistent participant was excluded because he conferred the highest utility to the worst probability.

Table 6 Number of participants who were willing to trade a loss[1] in their most important outcome for the largest gain[2] in another outcome

	Willingness to trade-off smallest loss in most important outcome against largest gain in other outcome			
Most important outcome N_{Max}[3]	Five-year survival N/N_{Total}[4] (%)	Five-year local control N/N_{Total} (%)	Faecal incontinence N/N_{Total} (%)	Sexual dysfunction N/N_{Total} (%)
Five-year survival (N = 7)	–	3/6 (50)	1/1 (100)	1/2 (50)
Five-year local control (N = 23)	16/18 (89)	–	8/9 (89)	4/7 (57)
Faecal incontinence (N = 30)	10/15 (67)	16/22 (73)	–	6/8 (75)
Sexual dysfunction (N = 6)	3/4 (75)	2/2 (100)	3/3 (100)	–

[1]Range of loss: 66 vs. 70% (probability of survival); 94 vs. 99% (probability of local control); 40 vs. 20% (risk of incontinence); 40 vs. 30% (risk of sexual dysfunction in men); or 30 vs. 10% (risk of sexual dysfunction in women).
[2]Range of benefit: 70 vs. 65% (probability of survival), 99 vs. 89% (probability of local control), 20 vs. 80% (risk of incontinence), 30 vs. 60% (risk of sexual dysfunction in men), and 10 vs. 70% (risk of sexual dysfunction in women).
[3]N_{Max} shows the total number of participants who considered that outcome as most important.
[4]N_{Total} represents the number of participants who were consistent in the rank order of estimated outcome-probability utilities, in the relevant pair of outcomes.

participant age for local control (Table 5). Consistency was positively related to importance for all outcomes, except for local control and for sexual dysfunction in female participants (Table 5). Consistency was not significantly related to education (data not shown).

Compensatory decision making

Respectively 39 (59%), 26 (39%) and 17 (26%) participants had consistent rank orders and were willing to trade the smallest decrease in survival for the largest improvement in local control, incontinence or sexual dysfunction. These numbers excluded three (5%) participants who had consistent rank orders and were unwilling to trade survival.

We assessed whether participants were willing to trade a loss from the best to the next-best probability (e.g. a decrease from 99% to 94% probability of local control or an increase from 20% to 40% probability of incontinence) in the outcome they considered most important for the largest benefit in any of the other outcomes. At least half of the participants were willing to trade such loss (Table 6). Overall, 58 (88%) participants were willing to trade a deterioration in their most important outcome for the largest

benefit in at least one other outcome, and were consistent in the respective rank orders.

ACA test and retest

Thirty-eight participants were asked to complete the ACA questionnaire again and 32 (84%) agreed. Two of them did not perform the retest because other commitments (1) or illness (1) did not allow a retest within a month. Two participants refused because they were critical about the task. One said it to be tedious, very annoying and showing only marginal differences between options. The other participant said she lost her way. Four participants refused because they considered it too burdensome (2), not useful (1) or because of time constraints (1). One participant was not approached for the retest as he was extremely negative after completing the ACA once. He said the exercise asked nearly moral questions, that differences in probabilities for outcomes were marginal and therefore made him start to think only superficially about them, and that the use of a computer made the exercise cold and impersonal.

Thirty participants performed the retest between 11 and 26 days after the interview (mean = 17, SD = 4.1). Mean importance for all outcomes at test and retest did not significantly differ (Table 7).

ACA- and TTM-based treatment preferences

Overall, 37/66 (56%) preferred surgery alone based on the ACA compared to 24 (36%) based on the TTM. This difference was significant (P = 0.019). Within individuals, ACA- and TTM-based preferences were identical for 39/66 (59%) participants.

ACA-based preference was not affected by having undergone PRT oneself (χ^2 = 1.7, NS). TTM-based preference was. Participants who had undergone PRT were more strongly in favour of PRT than participants who had not (χ^2 = 11.0, P = 0.001).

Table 7 Comparison of mean importance scores at test and retest (N = 30)

	Mean importance		
	Test	Retest	Test–retest[1]
Outcome	% ± SD	% ± SD	P-value
Five-year survival	21.0 ± 11.1	23.2 ± 9.8	0.24
Five-year local control	29.5 ± 8.9	29.6 ± 10.6	0.96
Faecal incontinence	27.5 ± 12.8	26.3 ± 11.9	0.63
Sexual dysfunction	21.9 ± 10.0	20.9 ± 9.8	0.48

[1]Paired t-test.

Discussion

To the best of our knowledge, this is the first study investigating ACA validity and reliability as a treatment outcome preference elicitation method and the first study to incorporate survival as an outcome. This study was conducted in a sample of disease-free rectal cancer patients. The treatment decision under consideration was preference for surgery vs. radiotherapy followed by surgery, a decision in which the participants had not been involved when they were diagnosed. Clinical results have been reported elsewhere.[37] In this paper, we focus on the methodological characteristics of ACA.

Almost all participants completed the ACA, and within a reasonable time. Two participants, who found it too difficult, did not complete the task. It is uncertain whether their difficulties referred to barriers on a cognitive, emotional or literacy level. Evaluating trade-offs in this context is difficult even to individuals who comprehend the task. It is therefore unclear whether they struggled with the trade-offs *per se*, or whether their difficulties lay with the task itself. Some participants' comments suggest that making the trade-offs is doable but demanding. Visual representations of the numerical probabilities and stating both probabilities of occurrence and non-occurrence might facilitate participants' comprehension.

With dominant pairs, participants most often preferred the best option, suggesting that they understood the task correctly. Still, one-fifth (7/32, 22%) of participants chose the dominated combination, a higher proportion than Stiggelbout *et al.*[22] (4/45, 9%) found. A likely

explanation for the disparity is Stiggelbout *et al.*'s use of qualitative labels rather than numerical probabilities in defining levels of ACA dimensions. The limited evidence available currently makes it difficult to state whether we should view the numbers as striking or not. Yet, these results suggest that most participants understood the task of paired comparisons and were able to make valid choices.

Consistency in the rank ordering of utilities of outcome-probabilities was examined by evaluating whether utilities for better probabilities were higher than utilities for worse probabilities. In only a minority of participants, utilities across all four outcomes were consistently rank ordered. However, the number of inconsistencies was low on average and inconsistencies occurred mostly between consecutive probabilities and almost never between an outcome's best and worst probability. Additionally, in most participants the best probability had the highest estimated utility. Some participants' comments suggest that they may have considered differences between probabilities small in the context of the presented trade-offs. Possibly, participants did not consider probabilities as appreciably different. Participants may further appear inconsistent in one outcome because they were attending to the other outcomes offered in paired combinations, and were in fact making trade-offs for other outcomes. Indeed, and similarly as Stiggelbout *et al.*[22] found, utilities were rank ordered consistently more often in outcomes that drove preferences more strongly. These may be evaluated with more scrutiny. In contrast to their results, consistency of rank orders was unrelated to education. Our results suggest that the task is doable across age and education in an older adult population in this context.

The assumption that ACA elicits compensatory decision making was tested by inspecting participants' willingness to trade survival and their most important outcome. Between one-fourth and more than half of the participants were found willing to trade the smallest loss in probability of survival for the largest benefit in any of the other outcomes, depending on the outcome being inspected. As our study was conducted among long-term survivors, it remains an empirical question whether individuals who have recently been diagnosed with cancer and need to decide about an upcoming treatment are willing to trade survival for better quality of life. Moreover, willingness to trade one's most important outcome for benefit in at least one of the other outcomes was apparent in a large majority of participants. These findings, similar to those reported by Stiggelbout *et al.*[22] suggest compensatory decision making in most participants and support the validity of using conjoint tasks in assessing relative treatment outcome importance in oncology. Of note, the results should be interpreted keeping in mind that the analyses were carried out only in the participants whose outcome-probability utilities were consistent in each respective pair of outcomes.

Not all participants appeared to trade. Simon[38] argued that individuals tend to minimize cognitive efforts and time taken by *satisficing*, i.e. deciding upon the option that meets a minimum criterion, rather then trying to obtain maximum utility. These participants may have applied a simplifying heuristic, such as choosing the option which was best on the most important outcome regardless of values on the other outcomes, following a *Take the Best* heuristic.[37] As Lloyd[37] puts forward, it is not clear how one should determine the exact nature of someone's decision-making strategies, especially as such processes are almost certainly unavailable to conscious reflection. Another possible explanation for participants' failure to trade is the combination of these participants' true preferences and the range of probabilities offered, which may not have been sufficiently broad to induce trading. We restricted the range of outcome-probabilities to probabilities that are clinically feasible for two main reasons: (i) the results would then represent patient preferences for outcomes that are within the reach of clinical practice and (ii) the task would not be lengthened unduly, as the more outcome-probabilities are included, the more paired comparisons are necessary for reliable estimations of outcome-probability utilities.

ACA mean outcome importance did not differ between test and retest, suggesting that it captures mean importance reliably over time.

Finally, ACA- and TTM-based treatment preferences differed in almost half of the participants. ACA-based treatment preferences were unrelated to past treatment, whereas TTM-based preferences strongly were. Higher preference for treatment modality with which patients have experience was found in other studies within oncology.[39–41] In these studies, preference elicitation techniques were used in which treatment modality was identified, as in the TTM. The ACA questionnaire did not identify treatment, thus modality preference could not induce biases in trade-offs. Moreover, unlike the TTM, ACA responses indicate to what extent treatment aspects drive patient preferences, given the probability ranges that were specified. This information is relevant and can be extrapolated to treatment options with similar key aspects but other probabilities.

Limitations

First, ACA is a 'main effects only' model, i.e. it assumes that there are no interactions among dimensions. It is questionable whether this assumption holds in our study. The validity of ACA findings may be threatened when dimensions are perceived to imply one another; these may be 'counted double' and thus be biased upwards in importance.[11] Clearly, this should be tested by comparing ACA with choice-based conjoint analysis results, a methodology which allows for interactions between dimensions.

Second, we used a labelled version of the TTM because it is usually applied as such. Moreover, all participants had received a letter with the results of the trial in which they had participated. So even without labels, a significant proportion of participants would probably have traced which description corresponded to which treatment. Nevertheless, it might be useful to compare the ACA to an unlabelled comparison.

Third, the exact probabilities for sexual dysfunction as presented in the TTM were not included in the ACA questionnaire. As we computed ACA-based preference on approximations of these probabilities, this may have affected the number of participants who had similar preferences when comparing the two assessment methods. However, of the three outcomes local control, incontinence and sexual dysfunction, the latter was least important in determining preference in half (35/66, 53%) of the participants and most important in only eight (12%). The use of approximations should therefore affect results only to a limited extent.

Finally, a select patient group was included. All participants had been treated for rectal cancer and all were long-term disease-free survivors. It is therefore questionable whether preferences may be generalizable to other cancer patient groups and, importantly, to patients who are still facing treatment.

Conclusion

Overall, the findings in disease-free rectal cancer patients suggest that in an oncology context, most individuals understand the ACA questionnaire, can complete it in a reasonable amount of time and most conform to the assumption of compensatory decision making. Also, in these long-term survivors, the exercise assessed mean outcome importance in a stable manner over time. ACA methodology therefore seems a valid treatment outcome preference elicitation method that captures preferences reliably over time. Treatment preferences derived from the holistic TTM and the decomposed ACA methodologies differed greatly. Different methods ask participants to perform different tasks, and may as a result invoke entirely different cognitive processes.[42] Which method should be preferred will thus strongly depend on the specific context and decision.

Conflicts of interest

The authors have no conflicts of interest.

Source of funding

Financial support was provided by a grant from the Dutch Cancer Society (grant number UL

2005-3213), Amsterdam, The Netherlands. Anne M. Stiggelbout was further supported by a VIDI award from the Netherlands Organization for Scientific Research NWO Innovational Research Incentives (grant number 917.56.356), The Hague, The Netherlands.

Acknowledgements

The authors thank all the patients for participating in this follow-up study. We are indebted to Petra Jellema and Frank Berkers for their contribution to the design of the ACA task, and to Monique Baas-Thijssen and Natasja Raaijmakers for their help in gathering the data. Also, we are grateful to Liana Fraenkel for her comments on an earlier draft of this paper.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix A Pairs of outcome-probabilities in which inconsistencies between rank orders of utilities occurred

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

1 Sepucha K, Ozanne E, Mulley AG Jr. Doing the right thing: systems support for decision quality in cancer care. *Annals of Behavioral Medicine*, 2006; **32:** 172–178.

2 Hlatky MA. Patient preferences and clinical guidelines. *JAMA*, 1995; **273:** 1219–1220.

3 Brothers TE, Cox MH, Robison JG, Elliott BM, Nietert P. Prospective decision analysis modeling indicates that clinical decisions in vascular surgery often fail to maximize patient expected utility. *The Journal of Surgical Research*, 2004; **120:** 278–287.

4 Cotler SJ, Patil R, McNutt RA *et al*. Patients' values for health states associated with hepatitis C and physicians' estimates of those values. *The American Journal of Gastroenterology*, 2001; **96:** 2730–2736.

5 Montgomery AA, Fahey T. How do patients' treatment preferences compare with those of clinicians? *Quality in Health Care*, 2001; **10** (Suppl. 1)**:** i39–i43.

6 Stalmeier PFM, Tol-Geerdink JJ, van Lin ENJT *et al*. Doctors' and patients' preferences for participation and treatment in curative prostate cancer radiotherapy. *Journal of Clinical Oncology*, 2007; **25:** 3096–3100.

7 O'Connor AM, Legare F, Stacey D. Risk communication in practice: the contribution of decision aids. *British Medical Journal*, 2003; **327:** 736–740.

8 Slovic P. The construction of preference. *American Psychologist*, 1995; **50:** 364–371.

9 Fraenkel L, Bodardus S, Wittnik DR. Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. *Medical Care*, 2001; **39:** 1203–1216.

10 Gan TJ, Lubarsky DA, Flood EM *et al*. Patient preferences for acute pain treatment. *British Journal of Anaesthesia*, 2004; **92:** 681–688.

11 Sawtooth Software. ACA 5.0 Technical paper. Sawtooth Software Technical paper series 2005. Available at: http://www.sawtoothsoftware.com, accessed 6 October 2008.

12 Tol-Geerdink JJ, Stalmeier PF, van Lin EN *et al*. Do patients with localized prostate cancer treatment really want more aggressive treatment? *Journal of Clinical Oncology*, 2006; **24:** 4581–4586.

13 Whelan T, Levine M, Willan A *et al*. Effect of a decision aid on knowledge and treatment decision making for breast cancer surgery: a randomized trial. *JAMA*, 2004; **292:** 435–441.

14 Flood AB, Wennberg JE, Nease RF Jr, Fowler FJ Jr, Ding J, Hynes LM. The Importance of patient preference in the decision to screen for prostate cancer: Prostate Patient Outcomes Research Team. *Journal of General Internal Medicine*, 1996; **11:** 342–349.

15 Onel E, Hamond C, Wasson JH *et al*. Assessment of the feasibility and impact of shared decision making in prostate cancer. *Urology*, 1998; **51:** 63–66.

16 Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ*, 2000; **320:** 1530–1533.

17 Fraenkel L, Bogardus ST, Concato J, Felson DT, Wittink DR. Patient preferences for treatment of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, 2004; **63:** 1372–1378.

18 Fraenkel L, Gulanski B, Wittink D. Patient treatment preferences for osteoporosis. *Arthritis and Rheumatism*, 2006; **55:** 729–735.

19 Beusterien KM, Dziekan K, Schrader S *et al*. Patient preferences among third agent HIV medications: a US and German perspective. *AIDS Care*, 2007; **19:** 982–988.

20 Ahmed SF, Smith WA, Blamires C. Facilitating and understanding the family's choice of injection device

for growth hormone therapy by using conjoint analysis. *Archives of Disease in Childhood*, 2008; **93:** 110–114.

21 Beusterien KM, Dziekan K, Flood E, Harding G, Jordan JC. Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment. *Value in Health*, 2005; **8:** 453–461.

22 Stiggelbout AM, Vogel-Voogt E, Noordijk EM, Vliet Vlieland TP. Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting? *Quality of Life Research*, 2008; **17:** 641–649.

23 Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber A. The effect of search procedures on utility elicitations. *Medical Decision Making*, 1998; **18:** 76–83.

24 Brundage MD, Davidson JR, Mackillop WJ, Feldman-Stewart D, Groome P. Using a treatment-tradeoff method to elicit preferences for the treatment of locally advanced non-small-cell lung cancer. *Medical Decision Making*, 1998; **18:** 256–267.

25 Elit LM, Levine MN, Gafni A *et al*. Patients' preferences for therapy in advanced epithelial ovarian cancer: development, testing, and application of a bedside decision instrument. *Gynecologic Oncology*, 1996; **62:** 329–335.

26 Brundage MD, Davidson JR, Mackillop WJ. Trading treatment toxicity for survival in locally advanced non-small cell lung cancer. *Journal of Clinical Oncology*, 1997; **15:** 330–340.

27 Kiebert GM, Stiggelbout AM, Leer JW, Kievit J, De Haes HJ. Test–retest reliabilities of two treatment-preference instruments in measuring utilities. *Medical Decision Making*, 1993; **13:** 133–140.

28 Stanek EJ, Oates MB, McGhan WF, Denofrio D, Loh E. Preferences for treatment outcomes in patients with heart failure: symptoms versus survival. *Journal of Cardiac Failure*, 2000; **6:** 225–232.

29 Salkeld G, Solomon M, Short L, Ryan M, Ward JE. Evidence-based consumer choice: a case study in colorectal cancer screening. *Australian and New Zealand Journal of Public Health*, 2008; **27:** 449–455.

30 Johnson FR, Ozdemir S, Mansfield C *et al*. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology*, 2007; **133:** 769–779.

31 Kapiteijn E, Marijnen CA, Nagtegaal ID *et al*. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *New England Journal of Medicine*, 2001; **345:** 638–646.

32 Peeters KC, van de Velde CJ, Leer JW *et al*. Late side effects of short-course preoperative radiotherapy combined with total mesorectal excision for rectal cancer: increased bowel dysfunction in irradiated patients – a Dutch colorectal cancer group study. *Journal of Clinical Oncology*, 2005; **23:** 6199–6206.

33 Marijnen CA, van de Velde CJ, Putter H *et al*. Impact of short-term preoperative radiotherapy on health-related quality of life and sexual functioning in primary rectal cancer: report of a multicenter randomized trial. *Journal of Clinical Oncology*, 2005; **23:** 1847–1858.

34 Peeters KC, Marijnen CA, Nagtegaal ID *et al*. The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. *Annals of Surgery*, 2007; **246:** 693–701.

35 Sprangers MA, te VA, Aaronson NK. The construction and testing of the EORTC colorectal cancer-specific quality of life questionnaire module (QLQ-CR38). European Organization for Research and Treatment of Cancer Study Group on Quality of Life. *European Journal of Cancer*, 1999; **35:** 238–247.

36 Gigerenzer G. The psychology of good judgment: frequency formats and simple algorithms. *Medical Decision Making*, 1996; **16:** 273–280.

37 Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? *Health Economics*, 2003; **12:** 393–402.

38 Simon HA. Rational choice and the structure of environments. *Psychological Review*, 1956; **63:** 129–138.

39 Jansen SJ, Kievit J, Nooij MA *et al*. Patients' preferences for adjuvant chemotherapy in early-stage breast cancer: is treatment worthwhile? *British Journal of Cancer*, 2001; **84:** 1577–1585.

40 McQuellon RP, Muss HB, Hoffman SL, Russell G, Craven B, Yellen SB. Patient preferences for treatment of metastatic breast cancer: a study of women with early-stage breast cancer. *Journal of Clinical Oncology*, 1995; **13:** 858–868.

41 Stiggelbout AM, Kiebert GM, de Haes JC *et al*. Surveillance versus adjuvant chemotherapy in stage I non-seminomatous testicular cancer: a decision analysis. *European Journal of Cancer*, 1996; **32A:** 2267–2274.

42 Froberg DG, Kane RL. Methodology for measuring health-state preferences–II: scaling methods. *Journal of Clinical Epidemiology*, 1989; **42:** 459–471.