



OPEN ACCESS

On selecting quality indicators: preferences of patients with breast and colon cancers regarding hospital quality indicators

Benjamin H Salampessy ¹, Ward R Bijlsma,² Eric van der Hijden,¹ Xander Koolman,¹ France R M Portrait¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2019-009818>).

¹Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, Noord-Holland, The Netherlands

²Department of Healthcare Procurement, Menzis, Enschede, The Netherlands

Correspondence to

Benjamin H Salampessy, Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, Noord-Holland, The Netherlands; b.h.salampessij@vu.nl

Received 20 May 2019

Revised 10 November 2019

Accepted 21 November 2019

Published Online First

12 December 2019



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Salampessy BH, Bijlsma WR, van der Hijden E, et al. *BMJ Qual Saf* 2020;**29**:576–585.

ABSTRACT

Background There is an increasing number of quality indicators being reported publicly with aim to improve the transparency on hospital care quality. However, they are little used by patients. Knowledge on patients' preferences regarding quality may help to optimise the information presented to them.

Objective To measure the preferences of patients with breast and colon cancers regarding publicly reported quality indicators of Dutch hospital care.

Methods From the existing set of clinical quality indicators, participants of patient group discussions first assessed an indicator's suitability as choice information and then identified the most relevant ones. We used the final selection as attributes in two discrete choice experiments (DCEs). Questionnaires included choice vignettes as well as a direct ranking exercise, and were distributed among patient communities. Data were analysed using mixed logit models.

Results Based on the patient group discussions, 6 of 52 indicators (breast cancer) and 5 of 21 indicators (colon cancer) were selected as attributes. The questionnaire was completed by 84 (breast cancer) and 145 respondents (colon cancer). In the patient group discussions and in the DCEs, respondents valued outcome indicators as most important: those reflecting tumour residual (breast cancer) and failure to rescue (colon cancer). Probability analyses revealed a larger range in percentage change of choice probabilities for breast cancer (10.9%–69.9%) relative to colon cancer (7.9%–20.9%). Subgroup analyses showed few differences in preferences across ages and educational levels. DCE findings partly matched with those of direct ranking.

Conclusion Study findings show that patients focused on a subset of indicators when making their choice of hospital and that they valued outcome indicators the most. In addition, patients with breast cancer were more responsive to quality information than patients with colon cancer.

INTRODUCTION

In many countries, policy makers have responded to rising healthcare expenditures by introducing managed competition.¹ One requirement for

optimal competition is that health-care consumers (patients) and those who choose on their behalf (eg, health insurers) have access to information on quality of care when choosing health-care providers and insurance plans.² The public reporting of quality information has therefore become an important focus for health policy.³ Measurement has generally been focused on hospital quality while using Donabedian's framework of structure, process and outcome indicators.⁴ Many quality measurement programmes, aimed at improving transparency, have been implemented. This has resulted in a proliferation of publicly reported quality indicators that are often published via user-friendly online platforms (eg, Centers for Medicare & Medicaid Services' (CMS) hospital compare site⁵).

While often presented with numerous quality indicators, previous research has indicated that patients are often unaware of quality differences across health providers, are overwhelmed by the large amount of available information, and if they do use such information to inform their decisions, they do so selectively.^{6,7} One way to improve responsiveness of patients to quality information is to optimise the presented information to them by, for example, restricting the number of indicators while focusing on the most important ones. Gaining insights into which quality indicators from the currently reported set patients value the most is crucial as efforts aimed at tailoring information to patients' needs could be designed more effectively when considering their preferences.

As in USA, managed competition has been implemented in the Netherlands. The Dutch health system aims to (1) Stimulate effective price and quality competition. (2) Encourages patient's choice. (3) Allows the selective contracting of providers by health insurers. (4) Offers universal access allowing patients to use healthcare covered by the basic health insurance package across all hospitals.^{2,8} To stimulate patient's choice, a vast amount of quality indicators is presented to patients via various online platforms. Given that patients only use these indicators in a very limited fashion,^{6,7} these platforms may benefit from a more user-tailored presentation of information.

This study estimated the preferences of groups of patients with breast cancer and colon cancer regarding publicly reported quality indicators for Dutch hospital care in order to optimise quality information presented to patients.

METHODS

Survey design and administration

We used discrete choice experiments (DCEs) to estimate patients' preferences. In a DCE, respondents are

presented with several hypothetical scenarios (choice sets) consisting of two or more alternatives that are systematically constructed by varying attributes across given levels. In each choice set, respondents choose that which in their opinion, and in accordance with the random utility maximisation framework, yields the highest utility and this therefore reflects their latent preferences as captured by the utility function.⁹ In this study, we employed two DCEs, that is, one for breast cancer and one for colon cancer, and conducted all study phases for each DCE separately.

We focused on breast and colon cancers as they were highly relevant given our focus on patients' choice for hospitals. On the one hand, breast cancer and colon cancer are considered important for the Dutch population: these conditions ranked in the top three of all types of cancer in the Netherlands in terms of burden of disease in 2015 and accounted for 2.1% and 3.4%, respectively, of all deaths in the Netherlands in 2017.^{10,11} On the other hand, the selected indicator sets have several advantages over other sets. Hospital quality of care for breast and colon cancers has been measured longer than for most other conditions: the sets for breast and colon

Table 1 Attributes and levels

Attributes	Description	Levels*	Expected sign
Breast cancer			
1) Volume	Total number of treated patients diagnosed with breast cancer	100, 350 and 450 patients	+
2) Waiting time 1	Average waiting time between diagnosis and tumour resection	15 days and 25 days	–
3) Waiting time 2	Average waiting time between diagnosis and combination surgery	25 days and 45 days	+
4) Preserved breast contour	Share of patients with a tumour resection whose breast contour was preserved	60%, 80% and 90%	+
5) Combination surgery†	Share of patients with a tumour resection who received a combination surgery	10%, 30% and 60%	+
6) Tumour residual‡	Share of tumour resections for which the tumour resection margin was shown to be irradical	1% and 5%	–
Colon cancer			
1) Volume	Total number of treated patients diagnosed with colon cancer	50, 150 and 200 patients	+
2) Waiting time	Share of patients whose waiting time between diagnosis and tumour resection amounted to less than 5 weeks	60%, 80% and 90%	+
3) Tumour residual§	Share of tumour resections for which the tumour resection margin was shown to be radical	85% and 95%	+
4) Complications¶	Share of patients with a tumour resection for which surgery-related complications occurred (risk-adjusted)	10% and 20%	–
5) Failure to rescue** ††	Share of patients with a tumour resection for which failure to rescue occurred (risk-adjusted)	5% and 15%	–

*Across all attributes, the lowest level was used as reference level.

†Type of surgery in which tumour resection and breast contour reconstruction are performed within a single operating session.

‡Resections are classified as irradical when tumour cells have been observed in the circumferential resection margin. Non-tumour-free margins are associated with unfavourable outcomes such as higher tumour recurrence rates.³¹

§In radical resections, contrary to irradical resections, no tumour cells have been observed in the circumferential resection margin and are associated with favourable outcomes (ie, disease-free survival).³²

¶Adjusted for gender, body mass index, age, Charlson Comorbidity Score, American Society of Anaesthesiologists (ASA) classification, preoperative tumour complications, tumour location, extent of resection due to invasive tumour growth or metastases, pT stadium (ie, primary tumour site) and pM stadium (ie, presence of distant metastasis) of the TNM-classification of malignant tumours.³³

**The share of patients with surgery-related complications who died while in hospital or within 30 days following surgery.¹² Relative to hospitals with high failure to rescue rates, those with low rates are considered to be more successful in saving a patient's life when surgery-related complications have occurred.³⁴

††Adjusted for all the listed case-mix variables under ¶, and for the urgency of the resection (acute or elective).

cancers have been implemented in the first few years of the governmental measurement programme, that is, since 2008 and 2011, respectively. Consequently, various initiatives have focused on the selected conditions first to present quality information via online platforms.

Attributes were selected from the current indicator set (ie, as published publicly by the National Health Care Institute for reporting year 2016¹²) by means of patient group discussions using a stepwise manner (see online supplementary material 1 for a detailed description). In short, the session for breast cancer consisted of three participants who unanimously identified 6 of 52 indicators as being the most important. The five participants in the colon cancer session identified 8 of 21 indicators as being the most important in the first round and finally reduced this number to 5 during the plenary discussion. The final subsets were included as attributes, while corresponding levels were based on the actual distribution in scores to reflect meaningful and realistic levels (table 1). To ensure that respondents thoroughly understood the attributes, levels and corresponding explanations, these texts were first checked by internal communication officers specialised in patient communication and then checked by patients (n=6) eligible as respondents. We used the original framing of the included indicators and determined the expected sign of coefficients based on literature. For a positively framed indicator, higher scores implied a higher quality level and thus a positive sign was expected, and vice versa.

For both DCEs, a D-efficient design (main effects only) was generated and blocked into three (breast cancer) and two (colon cancer) blocks each consisting of six choice sets per block. Each choice set consisted of two unlabelled alternatives (ie, hospitals). Given that hospital care for cancer is covered by the Dutch basic health insurance package and that patients only have to pay the small mandatory deductible amount when using covered healthcare,⁸ we hypothesised that, in real world settings, nearly all patients would at least choose a hospital, visit the physician and then decide whether or not to be treated. We decided therefore not to include an opt-out option as this would have made the choice set less realistic. An example of a given choice set included in each questionnaire is provided in online supplementary material 2. We piloted an initial DCE (experimental design for both conditions: D-efficient, main effects only, zero priors, two blocks with nine choice sets per block) among a small sample of the study population (breast cancer: n=15; colon cancer: n=10) and used the estimates to improve the final experimental designs. A priori sample size calculations (main effects only) based on Johnson and Orme's rule of thumb¹³ were estimated at 125 respondents in total for each DCE.

In the questionnaire we described the DCE task thoroughly, explained the attributes and levels, and provided an example of a choice set. Respondents were also asked to rank the attributes in order of importance in a single question (hereafter referred to as the *direct ranking*). In

addition, a *stated choice behaviour test* was included in which respondents chose a single statement that best described their behaviour ((1) Choose a hospital *independently*. (2) Use an online *comparative tool*. (3) *Be advised* by an expert (eg, general practitioners, family physicians)). The questionnaire also included questions related to personal characteristics (ie, age, gender and self-reported health), health literacy level (ie, the ability to obtain, process, and understand health information and services necessary to the making of decisions concerning health¹⁴) using the validated Dutch translation of Chew's 3-item Set of Brief Screening Questions¹⁵ and several open-ended feedback questions. After piloting, minor adjustments in wording were made but no changes to attributes and levels were necessary. The web-based questionnaires were distributed via patients' platforms between August and October 2017. Participation was anonymous and voluntary.

Econometric analysis

Analyses included choice data from the pilot and final questionnaire. Mixed logit models were used to account for the clustering of data (multiple choice sets per respondent) and for individual's preference variation,⁹ and modelled the following equations:

Equation 1 (breast cancer):

$$V_{ij} = (\beta_1 + \eta_{1i}) * \text{Volume (350patients)}_i + (\beta_2 + \eta_{2i}) * \text{Volume (450patients)}_i + (\beta_3 + \eta_{3i}) * \text{Waiting time1 (25days)}_i + (\beta_4 + \eta_{4i}) * \text{Waiting time2 (45days)}_i + (\beta_5 + \eta_{5i}) * \text{Preserved breast contour (80\%)}_i + (\beta_6 + \eta_{6i}) * \text{Preserved breast contour (90\%)}_i + (\beta_7 + \eta_{7i}) * \text{Combination surgery (30\%)}_i + (\beta_8 + \eta_{8i}) * \text{Combination surgery (60\%)}_i + (\beta_9 + \eta_{9i}) * \text{Tumor residual (5\%)}_i + \varepsilon_{ij}$$

Equation 2 (colon cancer):

$$V_{ij} = (\beta_1 + \eta_{1i}) * \text{Volume (150patients)}_i + (\beta_2 + \eta_{2i}) * \text{Volume (200patients)}_i + (\beta_3 + \eta_{3i}) * \text{Waiting time (80\%)}_i + (\beta_4 + \eta_{4i}) * \text{Waiting time (90\%)}_i + (\beta_5 + \eta_{5i}) * \text{Tumor residual (95\%)}_i + (\beta_6 + \eta_{6i}) * \text{Complications (20\%)}_i + (\beta_7 + \eta_{7i}) * \text{Failure to rescue (15\%)}_i + \varepsilon_{ij}$$

In each equation, the utility that individual 'i' derived from a hypothetical hospital alternative 'j' in a choice set was reflected by 'V_{ij}' and was characterised by the combination of levels on each attribute. The systematic part in 'V_{ij}' was reflected by the population's mean attribute utility weights ' β_1 - β_9 ' (equation 1) and ' β_1 - β_7 ' (equation 2), and the individual-specific variation in utility weights ' η_1 - η_9 ' (equation 1) and ' η_1 - η_7 ' (equation 2), while the random part was captured

by the error term ' ε_{ij} '. The latter was assumed to be normally distributed. Since we wanted to compare patients' responsiveness to quality indicators within and between the DCEs, attributes were included as dummy variables using a coding scheme in which the difference in actual scores between levels were standardised (see online supplementary material 3 for a detailed description). The interpretation of coefficients remained the same as that for standard dummy coding: for instance a change in *waiting time 1* from 15 (reference) to 25 days, *ceteris paribus*, was associated with a change in derived utility equal to the value of the estimated coefficient. In all analyses, we started with a full model (ie, random parameters for all attributes)

and changed random parameters to fixed parameters when corresponding SDs were not significant.

Probability analyses were conducted, as described by Lancsar *et al* (2007), to determine the relative impact of attributes.¹⁶ We first computed the probability of choosing an alternative reflected by only reference levels across all attributes (*base alternative*). We then changed one level of a given attribute to compute the percentage change in probability of choosing that specific alternative over the base alternative, thereby determining its relative impact. Since we used random parameters in our model we could not directly compute choice probabilities and therefore used simulations. The average choice probability was

Table 2 Characteristics of respondents

Patient group	Breast cancer		Colon cancer	
Sample	DCE sample (n=84)	Dutch patients with cancer*	DCE sample (n=145)	Dutch patients with cancer*
Demographics				
Gender (%)				
Male	1.2	0.6	26.2	53.2
Female	90.5	99.4	60.0	46.8
Missing	8.3	—	13.8	—
Age (%)				
18–39 years	10.7	4.9	9.6	1.0
40–59 years	69.0	37.1	31.8	11.0
60 years and older	19.1	58.0	55.9	88.0
Missing	1.2	—	2.7	—
Socioeconomic status				
Education level (%)				
Low	10.7	25.1	35.9	29.9
Moderate	38.1	42.6	31.7	39.4
High	48.8	32.3	25.5	30.7
Missing	2.4	—	6.9	—
Health				
Subjective health (%)				
(Very) poor	4.7		11.7	
Moderate	39.3		35.2	
(Very) good	53.6		49.0	
Missing	2.4		4.1	
Health literacy†				
Mean score (SD)	3.5 (0.5)		3.4 (0.6)	
Inadequate level	1.4		4.0	
Adequate level	98.6		96.0	
Stated choice behaviour‡				
Preferences respondent (%)				
Be advised by an expert	34.5		43.5	
Use a comparative tool	9.5		3.4	
Choose independently	54.8		26.9	
Missing	1.2		26.2	

*Source^{11 19}

†Measured only in the main study questionnaire using the validated Dutch translation of Chew's 3-item set of Brief Screening Questions.¹⁵

‡A single question in which respondents chose one statement that described their choice behaviour best.

DCE, discrete choice experiment.

Original research

computed by taking the average of all 5000 simulated probabilities.

In additional analyses, our main analyses were repeated using inverse probability weighting (IPW) in order to make our findings more representative to the Dutch patient populations with breast cancer and

colon cancers.¹⁷ Weights were computed using iterative proportional fitting aimed to ensure that the weighted marginal totals of our samples resembled those of the corresponding whole population (ie, gender (colon cancer only) and age) and large representative samples (ie, educational level).^{11 18 19} Moreover, as subgroup

Table 3 Results of mixed logit models and probability analyses

		Coefficients	SE*	P value	Relative impact (%)§
Breast cancer					
<i>Attributes†</i>					
1) Volume	100 patients (ref.)	—	—	—	—
	350 patients	0.208	0.120	0.08	10.3
	450 patients	0.232	0.083	0.01	11.5
2) Waiting time 1 (<i>diagnosis—tumour resection</i>)	15 days (ref.)	—	—	—	—
	25 days	−0.219	0.077	0.01	10.9
3) Waiting time 2 (<i>diagnosis—combination surgery</i>)	25 days (ref.)	—	—	—	—
	45 days	−0.524	0.123	<0.01	25.6
4) Preserved breast contour	60% (ref.)	—	—	—	—
	80%	0.279	0.131	0.03	13.8
	90%	0.434	0.138	<0.01	21.4
5) Combination surgery	10% (ref.)	—	—	—	—
	30%	0.379	0.184	0.04	18.7
	60%	0.855	0.184	<0.01	40.3
6) Tumour residual	1% (ref.)	—	—	—	—
	5%	−1.999	0.332	<0.01	69.9
<i>SD of random parameters</i>					
6) Tumour residual	5%	0.964	0.182	<0.01	
<i>Number of individuals</i>					
84					
<i>Number of observations</i>					
972					
<i>Log-likelihood</i>					
−194.763					
Colon cancer					
<i>Attributes‡</i>					
1) Volume	50 patients (ref.)	—	—	—	—
	150 patients	0.129	0.084	0.12	6.5
	200 patients	0.217	0.047	<0.01	10.8
2) Waiting time (<i>diagnosis—tumour resection</i>)	60% (ref.)	—	—	—	—
	80%	0.307	0.088	<0.01	15.2
	90%	0.147	0.072	0.04	7.3
3) Tumour residual	85% (ref.)	—	—	—	—
	95%	0.157	0.059	0.01	7.9
4) Complications	10% (ref.)	—	—	—	—
	20%	−0.172	0.048	<0.01	8.6
5) Failure to rescue	5% (ref.)	—	—	—	—
	15%	−0.439	0.071	<0.01	20.9
<i>SD of random parameters</i>					
5) Failure to rescue	15%	0.480	0.094	<0.01	
<i>Number of individuals</i>					
145					
<i>Number of observations</i>					
1386					
<i>Log-likelihood</i>					
−409.463					

Ref. reference category in the dummy coding for which the derived utility is standardised to zero.

*Reflects the bootstrapped SE

†Based on theory, signs of coefficients were expected to be positive for attributes 2 and 6, and negative for attributes 1 and 3–5 (shown in [table 1](#))

‡Based on theory, signs of coefficients were expected to be positive for attributes 1–3, and negative for attributes 4 and 5 (shown in [table 1](#))

§Expressed as the percentage change in choice probability relative to the base alternative (reference levels across all attributes).

Table 4 Results of direct ranking, ranked by percentage ranked first

Attributes	Mean rank	SD	Ranked first (%)	Ranked top 3 (%)*
Breast cancer†				
4) Preserved breast contour	5.10	1.69	73.8	83.8
1) Volume	3.60	1.59	10.8	53.8
5) Combination surgery	3.21	1.57	7.7	38.5
3) Waiting time 2	2.89	1.53	3.1	35.4
6) Tumour residual	2.98	1.24	3.1	21.5
2) Waiting time 1	3.22	1.34	1.5	47.7
Colon cancer‡				
5) Failure to rescue	3.67	1.31	33.9	67.2
3) Tumour residual	3.13	1.53	33.9	64.4
1) Volume	2.43	1.51	13.6	37.3
4) Complications	2.70	1.36	11.9	42.4
2) Waiting time	3.22	1.03	6.8	72.9

*Can add up to more than 100%.

†Scaled from 1 to 6 (from least to most important).

‡Scaled from 1 to 5 (from least to most important).

analyses we assessed whether preferences differed across subgroups by incorporating interaction terms between the attributes' levels and covariates (ie, gender (colon cancer only), age, educational level, subjective health and stated choice behaviour).

Experimental designs were generated in NGene V.1.1.2 (ChoiceMetrics, Australia). In mixed logit models, random parameters were assumed to be normally distributed and were estimated based on 5000 Halton draws. Model fit was assessed based on $-2\log$ likelihood functions. Models of main analyses were bootstrapped (5000 bootstraps with replacement²⁰). All models were estimated in Stata V.14.1 (StataCorp, College Station, Texas, USA). Results were considered statistically significant if $p < 0.05$.

RESULTS

Study population

The breast cancer questionnaire was completed by 84 respondents (table 2). Most respondents were female, aged between 40 years and 59 years, had attained a high educational level, perceived their health as good to very good, had an adequate health literacy level and stated that they preferred to choose a health provider independently. On average and compared with the breast cancer population, more respondents were female, relatively young and had attained a higher educational level.

The colon cancer questionnaire was completed by 145 respondents (table 2). Most of these respondents were female, aged 60 years or older, had attained a low education level, perceived their health as good to very good, had an adequate health literacy level and indicated that they preferred to be advised by an expert when choosing a health provider. Relative to the colon cancer population, the average respondent

was more likely to be male, younger and had attained a lower educational level.

Preferences

Regarding breast cancer, the mixed logit model demonstrated theoretical validity as all attributes had their expected direction (ie, sign) and had levels that were significant (table 3). The latter also implied that the sample had enough statistical power to detect main effects. Similar to the attribute selection phase in which patients unanimously identified 6 of 52 indicators as being the most important, the model indicated an obvious preference for specific attributes shown by a large difference in marginal utility across attributes; the largest coefficient was observed for *tumour residual* and the smallest coefficient for *waiting time 1* (coefficient (SE): -1.999 (0.332) and -0.219 (0.077), respectively). Preference heterogeneity was observed for *tumour residual*: as indicated by its negative sign for the mean coefficient, most respondents preferred low over high scores, while very few respondents (ie, 1.9%) preferred the opposite. The probability analysis (table 3) revealed that the relative impact was the largest for *tumour residual* (69.9%) and the smallest for *waiting time 1* (10.9%); this value for *tumour residual* implied that, relative to the base alternative, the change from the reference level to the 5%-level was estimated to a percentage change in choice probability equal to 69.9%. Hence, the probability analysis indicated that, on average and ceteris paribus, scores on *tumour residual* affected a respondent's choice of hospital the most. In the *direct ranking* (table 4), *preserved breast contour* was, on average, considered the most important attribute by respondents. Scaled from 1 to 6 (from least to most important), this attribute scored an average of 5.10 (SD: 1.69) and

was ranked first in 73.8% of all rankings. In addition, *volume* (53.8%) and *waiting time 1* (47.7%) were ranked frequently in respondents' top three.

Regarding colon cancer, the mixed logit model demonstrated theoretical validity as attributes had their expected direction and had levels that were significant (table 3). The presence of attributes with significant levels also indicated that the sample was large enough to detect main effects. Relative to the model for breast cancer and in line with the attribute selection process for colon cancer in which patients identified initially 8 and finally 5 of 21 indicators as the most important ones, differences in marginal utility in terms of absolute magnitude across attributes were smaller; the largest coefficient was observed for *failure to rescue* and the smallest coefficient for *tumour residual* (mean (SE): -0.439 (0.071) and -0.157 (0.059), respectively). For *failure to rescue*, the model indicated that most respondents preferred low over high scores, while 18.0% of the respondents preferred the opposite. In the probability analysis (table 3), the relative impact was the largest for *failure to rescue* (20.9%) and the smallest for *tumour residual* (7.9%); for *failure to rescue*, this value indicated that, relative to the base alternative, the change from the reference level to the 15%-level corresponded to an estimated percentage change in choice probability equal to 20.9%. The probability analysis (table 4) revealed that, on average and ceteris paribus, a respondent's choice of hospital was most affected by *failure to rescue*. When attributes were *directly ranked* from 1 to 5 (from least to most important), *failure to rescue* was, on average, ranked as the most important attribute by respondents (mean (SD) 3.67 (1.31)). *Failure to rescue* and *tumour residual* were both ranked first in 33.9% of all rankings, with the former being ranked more frequently in the average top three, that is, 67.2% and 64.4%, respectively.

Additional analyses

In the interest of brevity, results of additional analyses are included in online supplementary material 4. In IPW analyses, preferences for representative samples (breast and colon cancers) in terms of gender, age and educational level remained similar. The relative importance of attributes (ie, the order of attributes in terms of relative impact) remained very similar with no meaningful differences. Moreover, no differences in preferences across subgroups were observed except for seven cases. Patients with breast cancer aged 60 years and older considered *volume* (350 patients), *volume* (450 patients), *combination surgery* and *tumour residual* as less important compared with their counterparts: the utility derived from the attribute decreased, meaning it became less important (mean (SE) main effects and interaction terms: 0.389 (0.031) and -0.485 (0.196); 0.333 (0.080) and -0.247 (0.074); 0.937 (0.140) and -0.252 (0.104); -2.350 (0.228) and 1.298 (0.093),

respectively). A similar attenuation in preferences was observed for patients with colon cancer aged 60 years and older regarding *tumour residual* (mean (SE) main effects and interaction term: 0.300 (0.085) and -0.275 (0.117)). Higher educated patients with colon cancer had, relative to those who were lower educated, a stronger preference for *waiting time*, *complications* and *failure to rescue*: the derived utility from the attribute increased (mean (SE) main effects and interaction terms: -0.092 (0.130) and -0.472 (0.196); -0.065 (0.089) and -0.340 (0.139); -0.315 (0.126) and -0.399 (0.197), respectively).

DISCUSSION

Principal findings

Our study shows that patients focused on a subset of indicators when making their choice of hospital (findings of patient group discussions and DCEs) and that they valued the outcome indicators *tumour residual* (breast cancer) and *failure to rescue* (colon cancer) the most. Subgroup analyses indicated some heterogeneity in preferences across respondents' age (breast and colon cancer) and educational levels (colon cancer only). Moreover, the range in percentage change in choice probability was larger for breast cancer relative to colon cancer (range: 10.9%–69.9% and 7.9%–20.9%, respectively) indicating that the responsiveness to quality indicators differed between both study populations. In addition, DCE findings partly matched with those of direct ranking.

Possible explanations and comparison with other studies

Our first result indicating that patients regarded a subset of indicators as important supports previous findings.^{6 7} Similarly, the DCEs' results showing that outcome indicators were valued the most, are in line with those of a systematic review that shows that attributes in DCEs reflecting outcomes of cancer care are more important to patients than those reflecting processes and costs.²¹ We observed a difference in patients' responsiveness to quality indicators across study populations, a finding that might be explained by the differences in patient characteristics. Most patients with breast cancer preferred to seek quality information and choose a hospital independently, whereas most patients with colon cancer preferred to be advised by a general practitioner or family physician. Furthermore, similar to Louviere and Islam (2008), we observed a discrepancy between findings of the DCEs and those of the direct ranking; for example, *preserved breast contour* had only the fourth largest impact in the DCE, while being ranked as highly important in direct ranking. Louviere and Islam ascribe the discrepancy in relative importance of attributes between DCEs and direct ranking methods observed in their study to the presence of an explicit context described in the former that forces respondents to make a trade-off.²²

In our direct ranking, respondents ranked the attributes in order of importance while considering only their reflected quality domain. In contrast, in DCE the respondents also considered the realistic scores of an attribute (ie, levels), information regarding the actual distribution in scores (ie, included as accompanying information in the choice set) and were forced to make a trade-off across attributes. Hence, patients with breast cancer may value *preserved breast contour by itself* as highly important as indicated by direct ranking results, but in the explicit decision context of our choice set, they are willing to trade off preferable scores on this attribute in order to obtain more preferable scores on other attributes. Given that real life choices are also made in contexts and following Louviere and Islam's argumentation,²² our DCE findings may therefore be more likely to resemble the *true* relative importance of quality indicators compared with those estimated by our direct ranking.

Strengths and limitations

A strength of our study is that we focused solely on the perspective of patients. While using another perspective (eg, those of health providers) might result in a different relative importance of quality indicators, the patient perspective should take precedence over those of others when patients are considered to select their provider.²³

Moreover, we focused on currently reported quality indicators since we aimed to identify the most important ones in order to reduce the current set. Our goal was not to identify other relevant quality attributes that, according to patients, should be included as quality measures as such inclusion will expand the large current set. While policy makers are keen on using recently developed measures such as patient-reported outcome measures and patient-reported experience measures, so far only a few countries (eg, England, Sweden, parts of USA) have implemented these measures.^{24 25} Similarly, we did not change the original framing of indicators in order to make the DCE as realistic as possible. In the context of patient's choice, policy makers designing publicly reported quality indicators should be aware of possible framing effects as patients respond more strongly to negatively framed indicators than to those positively framed.²⁶

In addition, the indicator sets for breast and colon cancers rely on internationally conducted research. Hence, they cover similar quality domains as those covered by indicators used in most other countries such as USA (see, for example, the specification of indicators measured by CMS²⁷). We therefore believe that our findings are to a large extent generalisable to other health systems.

Furthermore, we modelled our DCEs within the commonly used random utility maximisation framework. Under this framework, (1) Respondents are assumed to act as rationally behaving consumers. (2)

An individual's preferences are considered to be fixed, well defined and decision-context invariant. (3) And the individual is fully aware of these preferences.^{28 29} In accordance with this framework, we assume that our estimated coefficients may vary when specific attributes are included or excluded, but that they should not differ across treatment phase.

We also need to acknowledge certain study limitations. IPW analyses indicated that a representative sample in terms of gender, age and educational level would have similar preferences as our sample of respondents. However, since we did not have any information on an individual's tumour stage of disease at diagnosis or current treatment phase, we were unable to assess whether preferences differed across such subgroups. On the one hand, the fact that most respondents perceived their health as good to very good may suggest that they have entered their follow-up phase. On the other hand, the patients' platforms distributing the questionnaires indicated that their members consisted of a mixture of patients recently diagnosed with cancer and of those who went through treatment several years ago. Moreover, our samples were large enough to detect main effects, but they may lack sufficient statistical power in our subgroup analyses which could explain the few observed differences in preferences across subgroups.

Implications for clinicians and policy makers

Our study shows that patients consider a subset of the publicly reported indicator set when making their choice of hospital, and that patients' responsiveness to quality indicators differs within the patient population (ie, across patient characteristics) and between the different patient populations (ie, stronger responsiveness was observed for patients with breast cancer relative to patients with colon cancer). These findings underline the importance of tailoring hospital quality information to what patients value as important. Offering comparative information will only enable those patients who are proactive and engaged to choose providers in accordance with their preferences, while additional efforts are called for to ensure that others also have the ability to make well-informed choices. In addition, in the context of shared decision making, physicians should have knowledge of patients' population preferences regarding hospital quality indicators because they should use these preferences to guide patients who are unable to indicate their own preferences, or who prefer to be advised. Therefore, our findings will also help with the design of decision aids, specifically, the presented option grids: a brief comparison of treatments based on important attributes, such as quality of care, that matter to patients when making decisions (see Elwyn *et al*³⁰ for an example regarding breast cancer).

CONCLUSIONS

Our study shows that patients focus on a subset of quality indicators when making their choice of

hospital, and that they value outcome indicators the most. Our study also suggests that the responsiveness to quality information differs both within and between populations. The findings underline that tailoring the presented quality indicators to patients' preferences may optimise their use.

Acknowledgements The authors would like to thank all respondents who completed the questionnaire and Koen Ligtenberg for his help during the project. The authors would also like to thank the (anonymous) reviewers and editors for their constructive and highly useful feedback. Their input helped to improve the overall quality of the manuscript.

Contributors WRB conducted the patient groups discussions. WRB and BHS were involved in selecting attributes and levels, constructing the experimental design, and designing, planning and distributing the questionnaire. BHS conducted the statistical analysis and wrote the manuscript. All authors reviewed, provided critical input and approved the final version.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Benjamin H Salampessy <http://orcid.org/0000-0002-7006-8559>

REFERENCES

- Gaynor M, Town RJ. Chapter Nine - Competition in Health Care Markets. In: Pauly MV, McGuire TG, Barros PP, eds. *Handbook of health economics*. Amsterdam: Elsevier, 2011; 2. 499–637.
- van de Ven WPM, Beck K, Buchner F, et al. Preconditions for efficiency and affordability in competitive healthcare markets: are they fulfilled in Belgium, Germany, Israel, the Netherlands and Switzerland? *Health Policy* 2013;109:226–45.
- Dranove D. Chapter Ten - Health Care Markets, Regulators, and Certifiers. In: Pauly MV, McGuire TG, Barros PP, eds. *Handbook of health economics*. Amsterdam: Elsevier, 2011; 2. 639–90.
- Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260:1743–8.
- Centers for Medicare & Medicaid Services (CMS). What is Hospital Compare? Available: <https://www.medicare.gov/hospitalcompare/about/what-is-HOS.html> [Accessed July 2018].
- Damman OC, Hendriks M, Rademakers J, et al. How do healthcare consumers process and evaluate comparative healthcare information? A qualitative study using cognitive interviews. *BMC Public Health* 2009;9:423.
- Faber M, Bosch M, Wollersheim H, et al. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. *Med Care* 2009;47:1–8.
- Enthoven AC, van de Ven WPM. Going Dutch--managed-competition health insurance in the Netherlands. *N Engl J Med* 2007;357:2421–3.
- Louviere JJ, Hensher DA, Swait JD. *Stated choice methods: analysis and applications*. Cambridge University Press, 2000.
- Volksgezondheidszorg.info. Ranglijst aandoeningen op basis van ziekte last (in DALY's) (translated: Ranking diseases based on burden of disease (in DALYs)). Available: <https://www.volksgezondheidszorg.info/ranglijst/ranglijst-aandoeningen-op-basis-van-ziekte last-dalys> [Accessed Oct 2018].
- Statistics Netherlands. StatLine. Available: <https://opendata.cbs.nl/statline/#/CBS/nl/> [Accessed May 2019].
- Zorginstituut Nederland (the Dutch National Health Care Institute). Ziekenhuizen en Zelfstandige Behandelcentra | Medisch specialistische zorg (translated: Hospitals and Independent Treatment Centers | Medical Specialist Care). Available: <https://www.zorginzicht.nl/Paginas/Home.aspx> [Accessed Feb 2018].
- Johnson R, Orme B. Getting the most from CBC. sawtooth software research paper, 2003. Available: <https://www.sawtoothsoftware.com/download/techpap/cbcmst.pdf> [Accessed Oct 2018].
- Hibbard JH, Peters E, Dixon A, et al. Consumer competencies and the use of comparative quality information: it isn't just about literacy. *Med Care Res Rev* 2007;64:379–94.
- Fransen MP, Stronks K, Essink-Bot ML. Gezondheidsvaardigheden: Stand van zaken (translated: health skills: state of affairs). *Laaggeletterdheid te lijf Den Haag Centrum voor ethiek en gezondheid* 2011;17:1–58.
- Lancsar E, Louviere J, Flynn T. Several methods to investigate relative attribute impact in stated preference experiments. *Soc Sci Med* 2007;64:1738–53.
- Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168:656–64.
- Kolenikov S. Calibrating survey data using iterative proportional fitting (raking). *Stata J* 2014;14:22–59.
- Aarts MJ, van der Aa MA, Coebergh JWW, et al. Reduction of socioeconomic inequality in cancer incidence in the South of the Netherlands during 1996–2008. *Eur J Cancer* 2010;46:2633–46.
- Efron B, Tibshirani RJ. *An introduction to the bootstrap* Boca Raton. Florida CRC press, 1994.
- Bien DR, Danner M, Vennedey V, et al. Patients' preferences for outcome, process and cost attributes in cancer treatment: a systematic review of discrete choice experiments. *Patient* 2017;10:553–65.
- Louviere JJ, Islam T. A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. *J Bus Res* 2008;61:903–11.
- Wheeler SB, Spencer J, Rotter J. Toward value in health care: perspectives, priorities, and policy. *N C Med J* 2018;79:62–5.
- Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
- Mediquest. Ontwikkelingen PROMs in 2018 (translated: Developments of PROMs in 2018).. Available: <https://home.mediquest.nl/deze-proms-zijn-verplicht-in-2018/> [Accessed Mar 2019].

- 26 Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? results from a discrete choice experiment in screening for colorectal cancer. *Value Health* 2009;12:354–63.
- 27 Centers for Medicare & Medicaid Services (CMS). CMS specifications manual version 5.2b. Available: <https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228775749207> [Accessed Feb2018].
- 28 McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, ed. *Frontiers in Econometrics*. New York: Academic Press, 1974: 105–42.
- 29 Lancaster KJ. A new approach to consumer theory. *J Polit Econ* 1966;74:132–57.
- 30 Elwyn G, Lloyd A, Joseph-Williams N, *et al*. Option grids: shared decision making made easier. *Patient Educ Couns* 2013;90:207–12.
- 31 Morrow M, Van Zee KJ, Solin LJ, *et al*. Society of surgical Oncology-American Society for radiation Oncology-American Society of clinical oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma in situ. *Ann Surg Oncol* 2016;23:3801–10.
- 32 Belluco C, De Paoli A, Canzonieri V, *et al*. Long-term outcome of patients with complete pathologic response after neoadjuvant chemoradiation for cT3 rectal cancer: implications for local excision surgical strategies. *Ann Surg Oncol* 2011;18:3686–93.
- 33 Sobin LH, Gospodarowicz MK, Wittekind C, eds. *TNM classification of malignant tumours*. 7. New Jersey, USA: John Wiley & Sons, 2011.
- 34 Almoudaris AM, Burns EM, Mamidanna R, *et al*. Value of failure to rescue as a marker of the standard of care following reoperation for complications after colorectal resection. *Br J Surg* 2011;98:1775–83.