

Patients' Preferences for Outcome, Process and Cost Attributes in Cancer Treatment: A Systematic Review of Discrete Choice Experiments

Daniela R. Bien¹ · Marion Danner² · Vera Vennedey² · Daniele Civello² · Silvia M. Evers³ · Mickaël Hiligsmann³

Published online: 31 March 2017

© The Author(s) 2017. This article is an open access publication

Abstract

Introduction As several studies have been conducted to elicit patients' preferences for cancer treatment, it is important to provide an overview and synthesis of these studies. This study aimed to systematically review discrete choice experiments (DCEs) about patients' preferences for cancer treatment and assessed the relative importance of outcome, process and cost attributes.

Methods A systematic literature review was conducted using PubMed and EMBASE to identify all DCEs investigating patients' preferences for cancer treatment between January 2010 and April 2016. Data were extracted using a predefined extraction sheet, and a reporting quality assessment was applied to all studies. Attributes were classified into outcome, process and cost attributes, and their relative importance was assessed.

Results A total of 28 DCEs were identified. More than half of the studies (56%) received an aggregate score lower than 4 on the PREFS (Purpose, Respondents, Explanation,

Findings, Significance) 5-point scale. Most attributes were related to outcome (70%), followed by process (25%) and cost (5%). Outcome attributes were most often significant (81%), followed by process (73%) and cost (67%). The relative importance of outcome attributes was ranked highest in 82% of the cases where it was included, followed by cost (43%) and process (12%).

Conclusion This systematic review suggests that attributes related to cancer treatment outcomes are the most important for patients. Process and cost attributes were less often included in studies but were still (but less) important to patients in most studies. Clinicians and decision makers should be aware that attribute importance might be influenced by level selection for that attribute.

Key Points for Decision Makers

Outcome attributes regarding effectiveness and adverse effects are most often included within discrete choice experiments in cancer treatment, and are often considered the most important by patients.

Process and cost attributes, in contrast, are included less often but are still of importance in most studies.

Clinicians and decision makers should be aware that patients value not only the outcome but also process and cost attributes, and aligning care with the patients' preferences could lead to improved adherence to treatment and therefore greater efficiency.

Electronic supplementary material The online version of this article (doi:10.1007/s40271-017-0235-y) contains supplementary material, which is available to authorized users.

✉ Mickaël Hiligsmann
m.hiligsmann@maastrichtuniversity.nl

¹ Faculty of Health Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

² Institute of Health Economics and Clinical Epidemiology, University of Cologne, Cologne, Germany

³ Department of Health Services Research, School for Public Health and Primary Care (CAPHRI), Maastricht University, PO Box 6161, 6200 MD Maastricht, The Netherlands

1 Introduction

As the population ages [1], expenditures are increasing, particularly in oncology care [2], and the efficient allocation of scarce resources is a key challenge for both policy makers and healthcare professionals [3, 4]. Rising expenses in healthcare give greater importance to the evaluation of interventions, financing and service delivery, which together entail the valuation of healthcare and health outcomes [3, 5]. In Germany, for instance, a reform of the pharmaceutical market (AMNOG [Pharmaceutical Market Reorganisation Act]) has been introduced in order to manage the costs of pharmaceuticals [6]. At the same time, taking patient preferences into account is seen as increasingly important, as patients are the payers and consumers of health technologies and services [7]. Matching healthcare policy with patient preferences might lead to the improved effectiveness of healthcare interventions by, for example, improving the adoption of and adherence to clinical treatments and public health programmes. Furthermore, those preferences can be useful in designing and evaluating healthcare programmes [8].

In order to uncover patients' preferences, several choices have to be made with regard to the methods of preference elicitation. Overall, preference elicitation methods can be divided into 'revealed' and 'stated' preference methods [9]. By using surveys to elicit patient preferences for characteristics of hypothetical treatments in an experimental framework, stated preference methods enable the assessment the importance of attributes [9, 10]. In contrast, revealed patient preferences, which rely on observed data, are difficult to investigate and thus are rarely used in healthcare. Among stated preferences methods, discrete choice experiment (DCE), a specific form of conjoint analysis, has been used extensively to elicit preferences in healthcare [12]. A DCE is suitable for assessing the relative importance of attributes and levels, and for calculating trade-offs between them. The importance of attributes always depends on the other attributes included in a DCE and on the range of levels included for an attribute. The feasibility of the DCE method has already been investigated by several HTA agencies, including the German Institute for Quality and Efficiency in Health Care (IQWiG [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen]), and an approval decision supported by data from a DCE was taken by the US Food and Drug Administration (FDA) [11–13].

In DCEs, respondents are asked to make choices among hypothetical alternatives that are described by systematically varying attribute levels (e.g. extent of drug effectiveness, types of adverse effects or frequency of dosage) [9, 14, 15]. The identification and selection of attributes

and levels are fundamentally important to obtaining valid results, and a proper selection and descriptions are required [8, 15]. Attributes chosen to describe alternatives within DCEs can be categorised, overall, into three main categories: (1) outcome attributes such as effectiveness or adverse effects; (2) process attributes, such as the mode of administration or involvement in clinical decision-making; and (3) cost attributes [16]. It is known that health-related outcomes are important decision criteria for patients, clinicians, policy makers and payers in medical decision-making processes [17, 18]. The importance of processes and costs in healthcare, however, are investigated less often [16].

Within this research, attention is focused on DCE studies investigating patients' preferences for cancer treatment, as early death and disability caused by cancer have the highest total economic burden worldwide [2]. The total economic cost of cancer is estimated to be €126 billion in the European Union [19]. Overall, cancer is a progressive disease that can affect every part of the body [20]; by 2012, the burden of cancer had risen to approximately 14 million new cases per year and 8.2 million deaths per year [21]. Previous reviews of DCEs in oncology have focused mainly on the methodology, such as experiment design, estimation procedures and validity of responses [22], on the treatment application [12], or on the preferences with regard to cancer screening [23, 24]. To our knowledge, none of these reviews specifically focused on cancer treatment or synthesized the importance of outcome, process and costs attributes. Therefore, this review was designed to systematically review DCEs eliciting preferences for alternative cancer treatments, assess their reporting quality using a checklist, and classify treatment attributes into outcome, process and cost attributes in order to assess their relative importance. Therefore, reviewing and synthesising the importance of levels and attributes in cancer treatments across studies can provide important insights into the relative importance of specific levels and attributes to patients, as well as about the importance of certain types of attributes (e.g. outcome-, process- or cost-related) relative to each other. Consideration of patients' values and preferences in clinical decision-making could indeed improve treatment adherence and satisfaction and finally lead to improved efficiency in cancer care [20].

2 Methodology

This research was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement [25]. For simplicity, the adopted PRISMA approach is summarised under four headings: eligibility criteria (Sect. 2.1), search strategy (Sect. 2.2),

study identification and selection (Sect. 2.3), and data extraction and quality appraisal (Sect. 2.4).

2.1 Eligibility Criteria

The methodology of this research builds on previous systematic reviews of DCEs in health economics by de Bekker-Grob et al. [22] and Clark et al. [12]. Only studies measuring stated preferences for cancer treatment using DCEs published between 2010 and April 2016 were included. A study was considered to be a cancer treatment study when the goal of the examined intervention was to cure or considerably prolong the life of patients or to ensure the best possible quality of life to cancer survivors [26]. DCEs were included when either cancer patients or parents of children with cancer participated. Studies using matching methods, multi-criteria decision methods, adaptive conjoint analyses or other preference methods were excluded, as were studies focusing on the methodology of stated preference methods and reviews of DCEs. Studies in languages other than English, German, Dutch or French were not taken into account.

2.2 Search Strategy

To include all relevant DCEs, a two-stage search process was conducted: (1) ‘reference searching’; and (2) ‘term searching’. The combination of different search strategies helped to ensure the internal validity of the review. In reference searching, references of earlier systematic reviews of DCEs [12, 22] were examined in order to identify those related to cancer treatment between 2010 and 2012. For term searching, the MEDLINE and EMBASE databases were screened to identify studies between 2010 and April 2016. Screening of both databases is beneficial in order to retrieve as many articles as possible, since EMBASE provides broader coverage of European journals [27]. To reduce the risk of missing relevant sources, a manual search within key journals and the reference lists of identified articles was also conducted [28].

Search terms were based on previous reviews of de Bekker-Grob et al. [22] and Clark et al. [12]. For MEDLINE searches, respective MeSH terms were used to ensure the inclusion of all synonyms of DCEs and additional terms related to cancer treatment; searches within EMBASE included the search function ‘EMTREE’. In addition to the terms, a combination of subject terms and free-text terms was included: ‘neoplasm’, ‘cancer’, ‘therapeutics’ and ‘treatment’. The full electronic search strategies for both MEDLINE and EMBASE are shown in Electronic Supplementary Material Appendix A.

2.3 Study Identification and Selection

For reliability purposes, the identification and selection of articles was carried out by two independent reviewers and contained three stages of screening. The first stage included screening of titles and abstracts and was based on the predefined eligibility criteria. In the second stage, selected articles were screened based on the full text. The third stage consisted of manual searches of the reference list of identified articles. In case of discrepancies, consensus was reached with the help of a third researcher.

2.4 Data Extraction and Reporting Quality Assessment

Data extraction and reporting quality assessment were performed in four steps. First, DCEs were systematically reviewed and study characteristics were summarised in a spreadsheet. Extracted characteristics included title, author, year of publication, country, study objective, population, cancer type and sample size. Second, assessment of the reporting quality was carried out by two independent reviewers (DRB and MH or MD) using a checklist. This checklist merged several items of the ISPOR (International Society For Pharmacoeconomics and Outcomes Research) checklist [8], which is an extensive guideline for setting up good practice conjoint analyses, with items of the 5-item PREFS (Purpose, Respondents, Explanation, Findings, Significance) checklist by Joy and colleagues [29]. Items from the PREFS checklist included a well-defined research question (purpose), an appropriate data collection instrument (respondent information), methods explained in sufficient detail (explanation), valid results and conclusions (findings), and appropriate statistical analyses (significance) [8, 29]. Each item received a binary score (acceptable or unacceptable) and an aggregate sum score (ranging from 0 to 5) was calculated and compared across studies in order to critically assess the reporting quality. In case of discrepancies, agreement between researchers was achieved. In addition, data reporting assessment based on the ISPOR checklist focused on items regarding ‘current practice’ in DCE development such as attribute and level identification, attribute and level selection, labelling, the average number of attributes and levels in the experiment, and the mode of survey administration [29].

In the third step, attributes were classified into three classes (outcome, process and cost) and subclasses (e.g. for outcome: effectiveness, adverse effects, quality of life). In case of doubt about the allocation of specific attributes, a second researcher was consulted. In the fourth and last step, we estimated the frequency of each attribute, then assessed if the attribute was significant and finally identified the most important attribute in each study. A

categorical attribute was considered significant if at least one coefficient level was significant (at a 5% level). When the relative importance was directly available, it was taken from the study after checking the calculations for correctness. If the relative importance was unavailable, but coefficients for attribute levels were provided, their relative attribute importance was calculated using the range method as recently proposed by the ISPOR Conjoint Analysis Good Research Practices Task Force [30]. Within this method, the range of attribute-specific levels is calculated by measuring the difference between the highest and lowest coefficient for the levels of the respective attribute. The relative importance is then calculated by dividing the attribute-specific level range by the sum of all attribute level ranges. The relative attribute importance calculated with this method always depends on the levels chosen and on the other attributes included in the experiment. In case of studies providing neither relative importance nor coefficients, the first author was contacted to get either relative importance or coefficients. If no data were provided, studies were excluded from the relative importance analysis. As an often used attribute had a higher chance of being selected as the most important, we compared the mean relative importance of the most important attributes per attribute class (outcome, process and cost) across the studies.

3 Results

Figure 1 provides an overview of the results of the study selection process. In total, 832 possible records were identified, from which 62 duplicates were removed. After title/abstract screening of the remaining 770 records, 718 studies were excluded. In the full-text assessment, another 26 articles were excluded, of which eight were published before 2010 and already included in the reviews by de Clark et al. [12] and de Bekker-Grob et al. [22]. Two additional articles were identified through manual searches of identified articles and included, resulting in 28 articles for the analysis.

3.1 Study Characteristics

An overview of the study characteristics is provided in Table 1. Within the investigated timeframe of this review (January 2010–April 2016), most studies were published in 2012 (25%) [31–37]. The majority of studies were conducted in North America ($n = 13$) [33, 34, 38–47] with at least one study each year since 2010; 2014 was a peak year, with five published studies. Europe has the second most publications ($n = 11$) [31, 35, 48–56], while Asia and

Australia, with two published studies each, make less common use of DCEs in cancer treatment [32, 36, 57, 58].

Sample sizes ranged from 89 up to 1096 participants, with an average of 272 participants per study (Table 2). While all studies gathered preferences of patients, five also included carers' preferences or those of healthcare providers [32, 36, 37, 50, 55]. A majority of the 25 studies (89%) targeted adult patients [31, 32, 34–36, 38–58], and only two included children or their parents [33, 37]. Furthermore, two studies focused their research on follow-up after therapy [49, 58] and one study on psychological care during the treatment process [32].

3.2 Reporting Quality Assessment

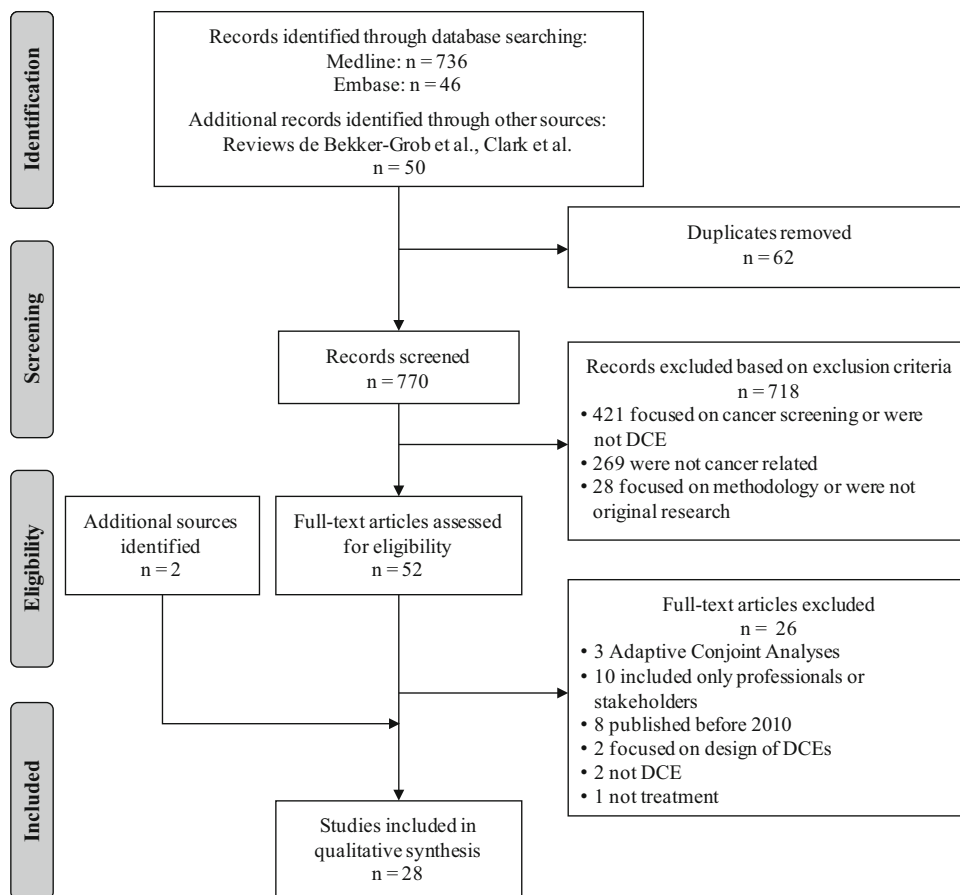
Table 2 includes the reporting quality assessment results, which are summarised by each letter of the acronym PREFS. All studies had a well-defined research question and clearly stated the purpose of the study in relation to preferences. However, only one study reported on the differences between responders and non-responders [52]. Methods of assessing preferences were clearly explained in all studies (96%), with one exception [50]. Just over half of the studies (52%) included all respondents in the analysis who at least partially completed the preference questions, or found that those excluded from analysis did not differ significantly from those included and were therefore found to have valid results and conclusions. Of 28 studies, 27 used significance tests to assess preference results. More than half of the studies (56%) received an aggregate score of 3 or less, one of which achieved a score of 2.

3.3 Current Practice in Data Generation

3.3.1 Attribute and Level Identification and Selection

A total of 93% of the studies reported methods for attribute identification [31, 32, 34–40, 42–58]; only two studies (7%) did not [33, 41]. Findings regarding current practice in attribute and level generation are summarised in Electronic Supplementary Material Appendix B. To identify attributes, most studies relied on a literature review of existing studies or package inserts (79%) [31, 32, 35–40, 42, 44, 46–52, 54–56, 58], followed by the use of qualitative research. In particular, expert interviews were a common method for attribute selection (57%) [32, 34, 36–40, 42, 43, 45, 47, 49, 50, 52, 55, 56], while patient interviews were conducted in 30% of studies [32, 38, 40, 42, 45, 46, 48, 52, 56, 57]. Decisions made by the research team or the specific use of patient focus groups were rare (7%) [32, 40]. Two studies used a pilot study in order to validate their identified attributes [42, 50].

Fig. 1 Flow diagram of study selection adapted from Moher et al. [25]. DCE discrete choice experiment



Within level identification, a literature review was the most commonly used method (44%), followed by expert interviews (33%). This is in contrast to attribute identification, where more studies ($n = 8$) did not report the identification of attributes (29%), but 20 studies did report level identification (71%).

3.3.2 Attribute Labelling and Number of Attributes and Levels

A total of 32% of the studies showed inconsistency in attribute labelling [33, 37–39, 48, 55], meaning that terms used for the results of analysis were different from those used within choice tasks and survey description (e.g. ‘severity of dehydration’ in the analysis in comparison with ‘need to seek additional treatment for dehydration’ in the choice task reference). Different attribute labelling might lead to biased results as participants may have a different understanding of the attributes when no consistent definition is provided. For example, one participant might already seek additional treatment for mild dehydration while another only seeks additional treatment for severe dehydration.

The number of included attributes is delineated in Table 1 and ranges mainly from 4 to 8, although two

studies included 11 attributes [46, 53]. Attributes included at minimum two levels, with a maximum of 5 levels.

3.3.3 Mode of Survey Administration

Nearly half of the studies (48%) used online questionnaires followed, in 37% of cases, by self-administered paper questionnaires. One study offered the possibility of choosing between an online or paper questionnaire [58]. Only five of the studies conducted face-to-face interviews (19%) [33, 36, 37, 48, 57], of which one was computer-assisted [48].

3.4 Classification of Attributes and Relative Importance

3.4.1 Attribute Classification

Of all the attributes included ($n = 168$) in the 28 studies, 118 were classified as outcome attributes (70%), followed by 41 process attributes (25%) and nine cost attributes (5%). The majority of studies (79%) combined the three attribute classes (outcome, process, cost) within the analysis ($n = 22$). Three studies included only outcome (11%)

Table 1 Study characteristics

Item	Category	All studies ($n = 28$) [n (%)]
Country of DCE ^a	Australia	2 (7)
	Canada	3 (11)
	France	2 (7)
	Germany	4 (14)
	Netherlands	4 (14)
	South Korea	1 (4)
	Spain	1 (4)
	Thailand	1 (4)
	UK	4 (14)
	USA	10 (36)
Year of publication	2010	1 (4)
	2011	4 (14)
	2012	7 (25)
	2013	3 (11)
	2014	6 (21)
	2015	5 (18)
	2016	2 (7)
Target population ^a	General population	1 (4)
	Children	1 (4)
	Parents	2 (7)
	Adults	25 (89)
	Elderly	9 (32)
	Physicians/healthcare provider	5 (18)
	Others ^a	2 (7)
Cancer type of interest	Breast	8 (30)
	Gastrointestinal stromal tumour	1 (4)
	Lung	2 (7)
	Lymphoma	1 (4)
	(Low-risk) basal cell carcinoma	2 (7)
	Oesophageal	2 (7)
	Ovarian	1 (4)
	Prostate	1 (4)
	Renal cell carcinoma	2 (7)
	Thyroid	1 (4)
N of attributes per study	Combination of different cancer types	7 (25)
	4	6 (21)
	5	6 (21)
	6	5 (18)
	7	4 (14)
	8	5 (18)
Average n of attributes	11	2 (7)
	Total	6.19
	Outcome	4.36
	Process	1.50
	Cost	0.32

DCE discrete choice experiment

^a More than one category per study possible

Table 2 Data extraction and quality appraisal of discrete choice experiment studies regarding cancer treatment

References	Sample size (n)	Attributes (n)	Outcome		Process			Costs			P	R	E	F	S	Score
			Effectiveness	Adverse effects	Quality of life	Mode of administration	Frequency of dosage	Waiting times	Location	Others						
Bridges et al. [31]	89	8	2	5	0	1	0	0	0	0	1	0	1	1	1	4
Goodall et al. [32]	152	6	0	0	0	0	0	0	0	6	0	1	0	1	1	4
Sung et al. [33]	274	4	0	3	0	0	1	0	0	0	1	0	1	0	1	3
Wong et al. [34]	272	8	1	6	0	1	0	0	0	0	1	0	1	1	1	4
Tinelli et al. [35]	174	4	2	1	0	0	0	0	0	0	1	1	0	1	1	4
Park et al. [36]	444	6	1	4	0	1	0	0	0	0	0	1	0	1	0	3
Regier et al. [37]	162	5	2	1	0	1	0	0	0	0	1	1	0	1	0	3
Hauber et al. [38]	173	8	0	8	0	0	0	0	0	0	0	1	0	1	1	4
Miller et al. [39]	301	8	0	7	0	0	0	0	0	0	1	1	0	1	0	3
Najafzadeh et al. [40]	1096	7	2	2	0	1	0	0	1	1	1	1	0	1	1	4
Smith et al. [41]	641	4	1	2	0	1	0	0	0	0	0	1	0	1	0	3
Johnson et al. [42]	296	4	0	2	0	0	1	0	0	0	1	1	0	1	1	4
Havrilesky et al. [43]	95	7	2	3	0	1	1	0	0	0	0	1	0	1	1	4
Lalla et al. [44]	298	8	0	7	0	0	0	0	0	0	1	1	0	1	0	2
Mohamed et al. [45]	138	7	1	6	0	0	0	0	0	0	0	1	0	1	0	3
daCosta DiBonaventura et al. [46]	181	11	1	8	1	1	0	0	0	0	0	1	0	1	0	3
Qian et al. [47]	580	6	2	2	0	1	0	0	0	0	1	1	0	1	0	3
Mühlbacher et al. [48]	211	7	2	4	0	1	0	0	0	0	0	1	0	1	0	3
Kimman et al. [49]	331	5	0	0	0	2	1	1	1	1	0	1	0	1	1	4
Thrumurthy et al. [50]	171	6	2	1	1	1	0	0	1	0	0	1	0	0	1	3
Hechmati et al. [51]	506	5	2	2	0	1	0	0	0	0	0	1	0	1	1	4
de Bekker-Grob et al. [52]	97	5	2	2	0	0	1	1	0	0	0	1	1	1	1	5
Martin et al. [53]	124	11	3	2	1	3	0	0	1	0	1	1	0	1	0	3
Mohamed et al. [54]	134	4	1	3	0	0	0	0	0	0	0	1	0	1	0	3
de Bekker-Grob et al. [55]	160	5	1	3	0	0	1	1	0	0	0	1	0	1	0	3
Damen et al. [56]	270	6	1	2	0	0	1	1	0	1	0	1	0	1	0	3
Ngorsuraches et al. [57]	146	4	1	2	0	0	0	0	0	0	1	1	0	1	0	3
Bessen et al. [58]	722	5	0	0	0	2	1	1	0	0	0	1	0	1	1	4

P purpose, R respondents, E explanation, F findings, S significance

[38, 45, 54] and three studies included only process attributes (11%) [32, 49, 58]. Electronic Supplementary Material Appendix C provides an overview of all classified attributes and corresponding studies.

Outcome ‘Progression-free survival’ was the most often included attribute regarding the effectiveness of treatment ($n = 8$) [31, 34, 36, 43, 45, 48, 54, 57], followed by ‘mortality’ ($n = 4$) [33, 37, 50, 52]. Most attributes were categorised regarding adverse effects (81 of 122 outcome attributes). ‘Nausea and vomiting’, ‘diarrhoea’ and ‘fatigue’ were the most commonly used adverse effects included in ten studies [31, 34, 37–39, 43–46, 48]. ‘Fever and infection’ as well as ‘hand–foot syndrome’ were also common ($n = 5$) [31, 37, 38, 42, 44]. ‘Quality of life’ was included in four studies [46, 50, 53, 56]; ‘cosmetic outcome’ was often used as a synonym for quality of life.

Process ‘Dosage form’ was by far the attribute most often included regarding the treatment process ($n = 11$) [31, 34, 36, 37, 41, 43, 46–48, 51, 53], followed by ‘frequency of dosage’ ($n = 8$) [33, 42, 43, 49, 52, 55, 56, 58] and ‘waiting time’ ($n = 2$) [49, 56].

Cost ‘Out-of-pocket cost’ was included in nine studies [35, 37, 39, 40, 42, 44, 47, 53, 57]. No other attributes regarding cost were included.

3.4.2 Significance of Attributes

It was possible to analyse attribute significance based on significant level coefficients in nearly all studies (86%). In three of four studies, attribute significance could not be analysed because no information was provided regarding p -values or confidence intervals [35, 44, 53]; in the fourth study the coefficients were not reported [41]. Figure 2 shows how often an attribute was significant and thus important to patients.

Of all 146 included attributes, 90% had significant level coefficients and were therefore of importance for patients. Furthermore, overall analysis of the significance shows that outcome attributes were most often significant, followed by process and cost attributes. Of 121 included outcome attributes, 95 were significant (79%). The coefficients for the levels of the outcome attributes related to adverse effects were more often significant than those for the outcome attributes effectiveness or quality of life. Of the coefficients for levels of process attributes, 71% were significant. The number of significant process attributes per subclass was more evenly distributed, with one peak (mode of administration). Six of nine cost attributes were significant (67%), i.e. important to patients.

3.4.3 Relative Importance of Attributes

Nearly 70% (19 of 28) of all studies either directly reported attribute importance based on the range method or reported

level coefficients, which are needed to calculate relative attribute importance using the range method. One study reported neither level coefficients nor attribute weights and was therefore excluded from the relative importance analysis [41]. Four studies provided level preference weight graphs but no values, making it impossible to calculate relative attribute importance [34, 38, 51, 54]. The authors of these studies were contacted, but only one replied whereby the relative importance could be calculated [54]. Five additional studies were excluded because the methods of calculating relative importance were not clearly stated and not reproducible [37, 46, 49, 53, 55]. The levels chosen for similar attributes (e.g. progression-free survival in months, certain adverse effects such as diarrhoea or fatigue) were very diverse across studies.

An overview of the relative importance scores per study is included in supplementary online appendix D. Figure 3 shows the number of times a class of attributes was the most important. The importance of an attribute was evaluated by comparing the relative importance scores of all attributes within a study; the attribute with the highest relative importance score was considered the most important.

The analysis shows that an outcome attribute was considered the most important in 14 of the 19 studies (effectiveness nine times, adverse effects four times, quality of life one time), followed by cost attributes in three studies and process attributes in two studies (mode of administration one time and others one time). The analysis further shows that process attributes were the most important only when no outcome or cost attributes were included. The second most important attributes were also mainly outcome related (14 times), followed by process-related ones (five times).

3.4.4 Mean Relative Importance

Comparison of the mean relative importance of an attribute class indicates that outcome attributes have the highest mean relative importance (37%) in comparison with process (33%) and cost (33%).

4 Discussion

This systematic review of DCEs focusing on cancer treatment preferences suggests that the number of DCEs has been constant over recent years (2010–2016), with an average of four publications per year. In contrast, the average number of all DCEs in healthcare rose continuously from a mean of three per year (1990–2000) to 14 (2001–2008) to 45 (2009–2012) [12, 22]. Most of the studies investigated the preferences of adults at risk of cancer or with treatment experience; however, there are

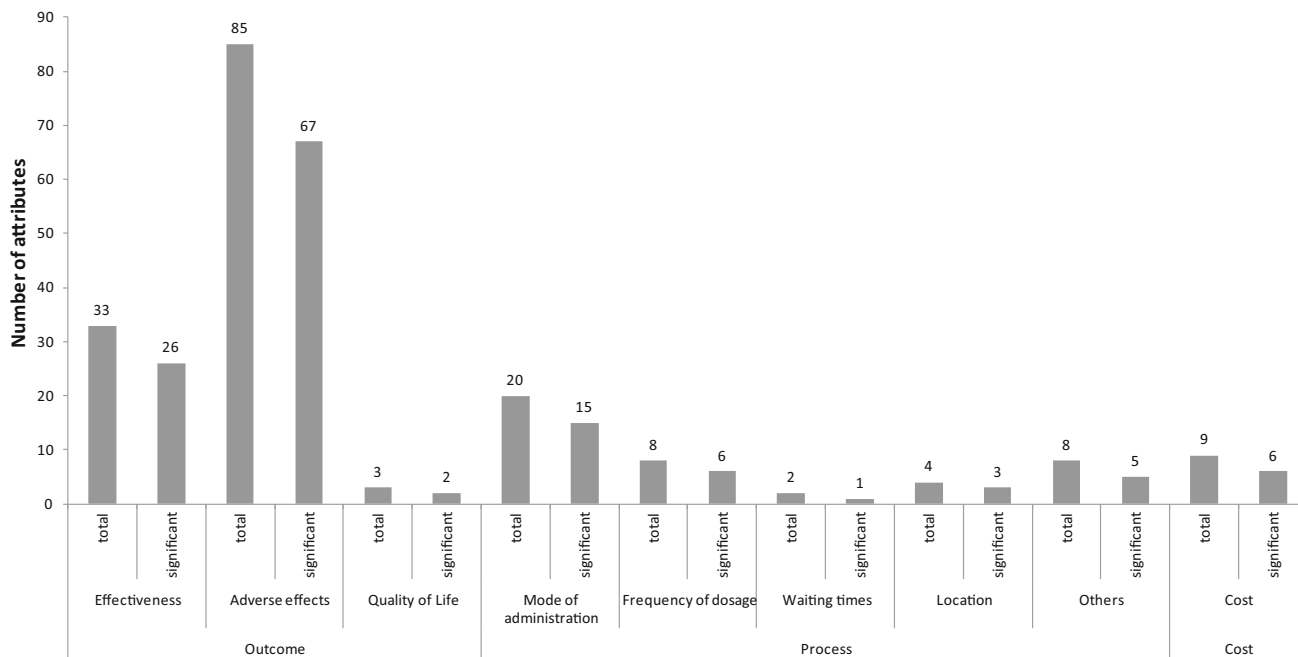


Fig. 2 Number of times an attribute class was used in all discrete choice experiments of cancer treatment and number of times an attribute class was significant; a categorical attribute was considered significant if at least one coefficient level was significant (at a 5% level)

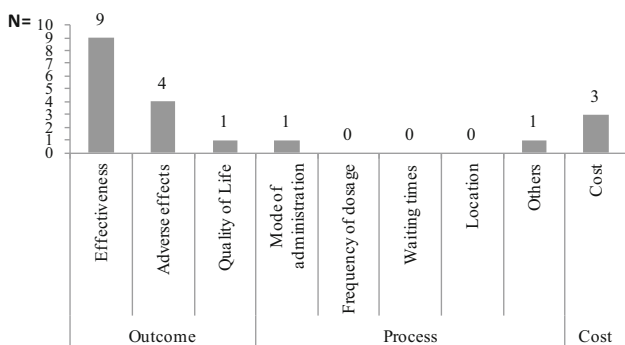


Fig. 3 Number of times an attribute class was the most important

also studies comparing patient’s preferences with those of healthcare providers or carers. For the purposes of our study, these latter studies have been analysed only regarding the preferences of patients.

Our review identified a number of shortcomings of current practice with regard to DCE applications in cancer treatment. First, despite growing awareness of the importance of patient involvement and research on patient preference, less than half of the cases (36%) carry out attribute identification and selection in cooperation with the patient. The majority of studies combined methods of attribute identification, reporting most often literature reviews (78%) and inclusion of professionals (71%). Excluding patients from attribute identification and selection might introduce omitted variable bias and might therefore lead to biased results on patients’ preferences,

since literature and professionals might deem different attributes important for inclusion in the analysis. Although we acknowledge that every research question does not require starting with patient input, it is often important to include patients at some point in time in attribute identification or selection, in refining attribute labelling and framing, or to pre-test the survey instrument on patients. Discrepancies between the preferences of professionals and patients have been reported by Thrumurthy et al. [50], who found differences in preferences between patients and doctors regarding mortality and quality of life. Also, Regier et al. [37] found that the importance ranking of effectiveness and adverse effect attributes differed between parents and healthcare professionals. Parents placed higher importance on the chance of infection, while healthcare professionals placed higher importance on chance of death [37]. Methods for preventing this type of bias are linked to the inclusion of patients in the process of attribute identification and selection and might be achieved through patient focus groups or patient interviews. In addition, pre-tests or pilot questionnaires could be used to increase the clarity of questions and the comprehensiveness of attributes relevant to patients.

A second potential issue is the administration of preference surveys. This review found that only 18% of the studies made use of an interviewer-led survey. To improve the data quality, the ISPOR guideline by Bridges et al. [8] suggests that interviewer-led administration of the survey may improve the quality of data because the interviewer

can recognize that more explanation is needed, can more fully explain the task and can answer questions (without leading the respondent). Interviewer-led administration, however, may suffer from limitations such as high financial and time expenses. Also, poorly trained interviewers might lead respondents or provide different information to different respondents, which could result in worse data quality than a self-administered survey. We therefore recommend that surveys be adequately piloted and assessed before use and that the mode of administration in each specific study setting and study population should be individually decided. Some populations might require more personal support than others. On the other hand, large samples are sometimes required and it might not be feasible to do interviewer-administered surveys.

Third, we did not assess the recruitment and sampling process of each study. It is, however, important that studies report on these processes in sufficient detail to allow for further analysis in case of a high proportion of non-response in a DCE. High non-response could, for example, indicate problems with getting access to the survey or with the survey itself (e.g. being too comprehensive, complicated, etc.).

Fourth, the number of included attributes varies across studies; this review found that there is no clear trend or standard regarding the amount of attributes to include, which is in line with other reviews [12]. Fifth, when looking at attribute labelling within DCEs, several inconsistencies have been highlighted. Terms used within the reporting vary partially in comparison with terms used within choice tasks. This inconsistency might cause problems with data interpretation based on published results.

Overall, the quality of reporting in DCEs was acceptable, but there is room for improvement. Even after the ISPOR guideline was implemented in 2011 [8], only 40% of studies achieved a score of 4 on the PREFS checklist and only one study achieved a score of 5 on the 5-point scale. For example, with one exception, all studies ranked an unacceptable score on the reporting of differences between responders and non-responders, which might lead to a non-response bias. Also, about half of the studies excluded some responders from the analysis but did not investigate the impact of these exclusions on study results. The most commonly noted reasons were that responders failed the comprehension test or did not answer enough choice tasks. Instead of excluding these respondents from the analyses, it would be interesting to understand why they, for example, failed comprehension questions or did not complete the survey, and to conduct sensitivity analysis on how their ex-/inclusion might impact study results. A respondent could have simply misinterpreted the test question but still have fully understood the remaining DCE questions.

Our review suggests that a similar number of studies included both outcome and process attributes; however, more attributes regarding treatment outcomes than treatment process were included. Our review suggests that most attribute levels (78%) are significant and thus important for patients. Based on the number of times that an attribute was said to be the most important, one may conclude that outcome attributes are the most important to patients. In this context one should keep in mind that attribute importance based on level ranges is always conditional on the levels chosen. Attribute importance may differ across studies, depending on the level selection of the same attribute in different studies [5]. For example, we included one study in metastatic breast cancer patients in this review that coded the adverse effect attribute 'hair loss' into the categories 0, 48 and 94% chance of losing most or all of your hair [46], while in another study in the same indication this attribute was coded as none/not noticeable hair loss versus obvious hair loss [44]. The large level range in the first study and the rather vague/weak level definition in the second one might have caused this attribute to rank as more important in the first study (hair loss ranked first before, for example, fatigue, ranked third, and diarrhoea, ranked sixth) and much less important in the second study (hair loss ranked sixth, fatigue ranked fifth and diarrhoea second). The second-ranking attribute diarrhoea in this study, on the other hand, was coded with levels from '2 stools or less per day' to an extreme level of 'being unable to leave the house' [44] compared with a much less extreme coding of levels in the first study [46] (i.e. having a 0, 5 or 15% chance of diarrhoea), where this attribute ranked only sixth. Also, including separate attributes for a specific type of adverse effect, its severity and frequency might generate much more detailed results than if one adverse effect attribute combines these dimensions in a single attribute, as can be seen by comparing the studies by Miller et al. [39] and Smith et al. [41]. Overall, the attribute and the attribute level choice across studies included in this review was very diverse and a comparison across studies was therefore difficult. Excluding/including specific attributes in a DCE further impacts the preference for the remaining attributes, which makes comparisons even more difficult to perform.

Our study showed that while an intervention's effectiveness and the prevention of adverse effects were usually most preferred by patients, process and cost attributes were also of, though of lesser, importance. Process and cost attributes may be less important relative to effectiveness attributes, but may still be important when it comes to real life (not only hypothetical) treatment choices or, for example, when it comes to decisions between treatment options where treatment effectiveness and adverse effects do not differ significantly.

Cost was included in only one-third of the studies, and especially in countries where health insurance is not obligatory (60%) and patients have a higher chance of paying costs out of pocket. Including a cost attribute in oncology studies could be challenging since the cost of cancer drugs could be extremely high. In addition to health insurance, there are financial aid programmes for patients through hospitals, drug companies and patient groups, and patients often end up paying very different amounts for the same drug. If patients can't afford any of the costs in the ranges offered in a DCE, then they could either ignore the attribute or dominate on cost. Although cost is certainly very important in treatment choice and adherence, including costs in a DCE as an attribute does not always work and more research could focus on how to estimate the impact of cost when costs are so high many patients can't afford them.

Although this review followed best practices in systematic reviews, it has some limitations. First, the PREFS checklist might not be comprehensive enough to assess all relevant aspects of (reporting) quality within DCEs. The checklist is limited to five questions and misses several important criteria that should be reported by DCE studies. In order to include additional important (reporting) quality aspects of DCEs in the analysis, we combined the PREFS checklist with aspects of the ISPOR checklist published by Bridges and colleagues [8]. While we assessed the reporting of data quality regarding these items, our study does not comprehensively address the potential limitations of study quality, which might, for example, have been caused by limited sample sizes or other methodological problems with data analysis.

Second, this review focuses on cancer treatment attributes; therefore, study results cannot be generalised to cancer screening or treatment of differing diseases. Third, the review could not include all studies in the analysis of relative importance. Based on the range method, calculation of relative importance could be reproduced in about 70% of the studies, including one that was calculated after the first author, upon request, provided the preference weights. It was apparent that various studies lacked transparency in the statistical analysis of DCE findings. Some of the studies where the range method was not applied used other methods to assess relative attribute importance, among them the comparative willingness-to-pay for attributes or just using the highest estimated level coefficient as the weight for that attribute. Further research on patients' preferences should explore the range of methods used in DCEs to assess attribute importance and their potential strengths and weaknesses. The range method proposed by the ISPOR Conjoint Analysis Good Research Practices Task Force may be the most often used, but it is not the only method to assess attribute importance [29]. Finally,

this review included only full articles already published and did not include conference protocols.

This research could provide valuable information for clinical and policy decision-making as it provides information about actual patient preferences. Involving patients in decision-making might improve adherence to medication and therefore lead to greater efficiency. Moreover, this review could guide further research in cancer preference studies, as it provides an overview and insight into the shortcomings of current practices in DCEs regarding cancer treatment.

5 Conclusion

In this systematic review of DCEs conducted to investigate patients' preferences for cancer treatment, we observed that outcome attributes regarding effectiveness and adverse effects are most often included within DCEs, and are often considered the most important by patients. Process and cost attributes, in contrast, are included less often but are still of importance in most studies. Clinicians and decision makers should be aware that patients value not only the outcome but also process and cost attributes, and aligning care with the patients' preferences could lead to improved adherence to treatment and therefore greater efficiency.

Acknowledgements Daniela R. Bien conducted the systematic review and drafted the manuscript. Mickaël Hilgsmann and Marion Danner supported data interpretation, reviewed the manuscript and facilitated the project. Vera Venedey and Silvia M. Evers critically reviewed and contributed to the draft manuscript. Daniele Civello supported the systematic search for literature.

Compliance with Ethical Standards

Conflict of interest D.R. Bien, M. Danner, V. Venedey, D. Civello, S.M. Evers and M. Hilgsmann declare that they have no conflicts of interest.

Funding No funding was used for the conduct of this study.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. European Commission. Demographic analysis. 2016. <http://ec.europa.eu/social/main.jsp?catId=502>. Accessed 31 Jan 2016.
2. John RM, Ross H. The global economic cost of cancer. *Livingstone/American Cancer Society*; 2010. <https://old.cancer.org/>

- acs/groups/content/@internationalaffairs/documents/document/acspc-026203.pdf. Accessed 17 March 2017.
3. Taylor R, Drummond M, Salkeld G, Sullivan S. Inclusion of cost effectiveness in licensing requirements of new drugs: the fourth hurdle. *BMJ*. 2004;329(7472):972–5.
 4. Sullivan M. The new subjective medicine: taking the patient's point of view on health care and health. *Soc Sci Med*. 2003;56(7):1595–604.
 5. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics*. 2008;26(8):661–77.
 6. GKV-Spitzenverband. AMNOG—evaluation of new pharmaceutical 2016 [updated 4 Apr 2016]. https://www.gkv-spitzenverband.de/english/statutory_health_insurance/amnog_evaluation_of_new_pharmaceutical/amnog_english.jsp. Accessed 24 Mar 2017.
 7. Institute for Quality and Efficiency in Health Care (IQWiG). Choice-based conjoint analysis—pilot project to identify, weight, and prioritize multiple attributes in the indication “hepatitis C”. 2014. https://www.iqwig.de/download/GA10-03_Executive-summary-of-working-paper-1.1_Conjoint-Analysis.pdf. Accessed 24 Mar 2017.
 8. Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health*. 2011;14(4):403–13.
 9. Carson RT, Louviere JJ. A common nomenclature for stated preference elicitation approaches. *Environ Res Econ*. 2011;49(4):539–59.
 10. Hiligsmann M, Bours SPG, Boonen A. A review of patient preferences for osteoporosis drug treatment. *Curr Rheumatol Rep*. 2015;17(9):61.
 11. Institute for Quality and Efficiency in Health Care (IQWiG). Allgemeine Methoden. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; 2014.
 12. Clark MD, Determan D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. *Pharmacoeconomics*. 2014;32(9):883–902.
 13. Ho MP, Gonzalez JM, Lerner HP, Neuland CY, Whang JM, McMurry-Heath M, et al. Incorporating patient-preference evidence into regulatory decision making. *Surg Endosc*. 2015;29(10):2984–93.
 14. Ryan M. Discrete choice experiments in health care. *BMJ*. 2004;328(7436):360–1.
 15. Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Econ Policy Law*. 2009;4(Pt 4):527–46.
 16. Schaarschmidt M-L, Schmieder A, Umar N, Terris D, Goebeler M, Goerdts S, et al. Patient preferences for psoriasis treatments: process characteristics can outweigh outcome attributes. *Arch Dermatol*. 2011;147(11):1285–94.
 17. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318(7182):527–30.
 18. Krousel-Wood MA. Practical considerations in the measurement of outcomes in healthcare. *Ochsner J*. 1999;1(4):187–94.
 19. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol*. 2013;14(12):1165–74.
 20. Hong WK, Bast JRC Jr, Hait WN, Kufe DW, Pollock RE, Weichselbaum RR, et al. Cardinal manifestations of cancer. In: Hong WK, Bast RC, Hait WN, Kufe DW, Pollock RE, Weichselbaum RR, Holland JF, Frei III E, editors. *Holland-Frei cancer medicine*. 8th ed. Shelton: People's Medical Publishing House-USA; 2010. p. 1–3.
 21. Press release No. 224. World cancer report 2014. Global battle against cancer won't be won with treatment alone: effective prevention measures urgently needed to prevent cancer crisis. Lyon: International Agency for Research on Cancer; 2014 Feb 4 [press release]. Lyon/London, 3 February 2014. https://www.iarc.fr/en/media-centre/pr/2014/pdfs/pr224_E.pdf. Accessed 18 March 2017.
 22. de Bekker-Grob E, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ*. 2012;21(2):145–72.
 23. Phillips KA, Van Bebber S, Marshall D, Walsh J, Thabane L. A review of studies examining stated preferences for cancer screening. *Prev Chronic Dis*. 2006;3(3):A75.
 24. Mansfield C, Tangka FKL, Ekwueme DU, Lee Smith J, Guy GP Jr, Li C, et al. Stated preference for cancer screening: a systematic review of the literature, 1990–2013. *Prev Chronic Dis*. 2016;13:E27.
 25. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9, W64.
 26. World Health Organization. Cancer: screening for various cancers. 2016. <http://www.who.int/cancer/detection/variouscancer/en/>. Accessed 24 Mar 2017.
 27. Wright RW, Brand RA, Dunn W, Spindler KP. How to write a systematic review. *Clin Orthop Relat Res*. 2007;455:23–9.
 28. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol*. 2011;11(1):15.
 29. Joy SM, Little E, Maruthur NM, Purnell TS, Bridges JFP. Patient preferences for the treatment of type 2 diabetes: a scoping review. *Pharmacoeconomics*. 2013;31(10):877–92.
 30. Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value Health*. 2016;19(4):300–15.
 31. Bridges JF, Mohamed AF, Finnern HW, Woehl A, Hauber AB. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer*. 2012;77(1):224–31.
 32. Goodall S, King M, Ewing J, Smith N, Kenny P. Preferences for support services among adolescents and young adults with cancer or a blood disorder: a discrete choice experiment. *Health Policy*. 2012;107(2):304–11.
 33. Sung L, Alibhai SM, Ethier M-C, Teuffel O, Cheng S, Fisman D, et al. Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. *J Clin Epidemiol*. 2012;65(6):627–34.
 34. Wong MK, Mohamed AF, Hauber AB, Yang J-C, Liu Z, Rogerio J, et al. Patients rank toxicity against progression free survival in second-line treatment of advanced renal cell carcinoma. *J Med Econ*. 2012;15(6):1139–48.
 35. Tinelli M, Ozolins M, Bath-Hextall F, Williams HC. What determines patient preferences for treating low risk basal cell carcinoma when comparing surgery vs imiquimod? A discrete choice experiment survey from the SINS trial. *BMC Dermatol*. 2012;12(1):1.
 36. Park M-H, Jo C, Bae EY, Lee E-K. A comparison of preferences of targeted therapy for metastatic renal cell carcinoma between the patient group and health care professional group in South Korea. *Value Health*. 2012;15(6):933–9.
 37. Regier DA, Diorio C, Ethier M-C, Alli A, Alexander S, Boydell KM, et al. Discrete choice experiment to evaluate factors that influence preferences for antibiotic prophylaxis in pediatric oncology. *PLoS One*. 2012;7(10):e47470.

38. Hauber AB, Gonzalez JM, Coombs J, Sirulnik A, Palacios D, Scherzer N. Patient preferences for reducing toxicities of treatments for gastrointestinal stromal tumor (GIST). *Patient Prefer Adher*. 2011;5:307–14.
39. Miller PJ, Balu S, Buchner D, Walker MS, Stepanski EJ, Schwartzberg LS. Willingness to pay to prevent chemotherapy induced nausea and vomiting among patients with breast, lung, or colorectal cancer. *J Med Econ*. 2013;16(10):1179–89.
40. Najafzadeh M, Johnston KM, Peacock SJ, Connors JM, Marra MA, Lynd LD, et al. Genomic testing to determine drug response: measuring preferences of the public and patients using Discrete Choice Experiment (DCE). *BMC Health Serv Res*. 2013;13:454.
41. Smith ML, White CB, Railey E, Sledge GW Jr. Examining and predicting drug preferences of patients with metastatic breast cancer: using conjoint analysis to examine attributes of paclitaxel and capecitabine. *Breast Cancer Res Treat*. 2014;145(1):83–9.
42. Johnson P, Bancroft T, Barron R, Legg J, Li X, Watson H, et al. Discrete choice experiment to estimate breast cancer patients' preferences and willingness to pay for prophylactic granulocyte colony-stimulating factors. *Value Health*. 2014;17(4):380–9.
43. Havrilesky LJ, Alvarez Secord A, Ehrisman JA, Berchuck A, Valea FA, Lee PS, et al. Patient preferences in advanced or recurrent ovarian cancer. *Cancer*. 2014;120(23):3651–9.
44. Lalla D, Carlton R, Santos E, Bramley T, D'Souza A. Willingness to pay to avoid metastatic breast cancer treatment side effects: results from a conjoint analysis. *Springerplus*. 2014;3:350.
45. Mohamed MAF, Hauber AB, Neary MP. Patient benefit-risk preferences for targeted agents in the treatment of renal cell carcinoma. *Pharmacoeconomics*. 2011;29(11):977–88.
46. daCosta DiBonaventura M, Copher R, Basurto E, Faria C, Lorenzo R. Patient preferences and treatment adherence among women diagnosed with metastatic breast cancer. *Am Health Drug Benefits*. 2014;7(7):386–96.
47. Qian Y, Arellano J, Hauber AB, Mohamed AF, Gonzalez JM, Hechmati G, et al. Patient, caregiver, and nurse preferences for treatments for bone metastases from solid tumors. *Patient*. 2016;9(4):323–33.
48. Mühlbacher AC, Bethge S. Patients' preferences: a discrete-choice experiment for treatment of non-small-cell lung cancer. *Eur J Health Econ*. 2015;16(6):657–70.
49. Kimman ML, Dellaert BG, Boersma LJ, Lambin P, Dirksen CD. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. *Acta Oncol*. 2010;49(3):328–37.
50. Thrumurthy S, Morris J, Mughal M, Ward J. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophago-gastric cancer. *Br J Surg*. 2011;98(8):1124–31.
51. Hechmati G, Hauber AB, Arellano J, Mohamed AF, Qian Y, Gatta F, et al. Patients' preferences for bone metastases treatments in France, Germany and the United Kingdom. *Support Care Cancer*. 2015;23(1):21–8.
52. de Bekker-Grob EW, Niers EJ, van Lanschot JJB, Steyerberg EW, Wijnhoven BP. Patients' preferences for surgical management of esophageal cancer: a discrete choice experiment. *World J Surg*. 2015;39(10):2492–9.
53. Martin I, Schaarschmidt ML, Glocker A, Herr R, Schmieder A, Goerdts S, et al. Patient preferences for treatment of basal cell carcinoma: importance of cure and cosmetic outcome. *Acta Derm Venereol*. 2016;96(3):355–60.
54. Mohamed AF, González JM, Fairchild A. Patient benefit-risk tradeoffs for radioactive iodine-refractory differentiated thyroid cancer treatments. *J Thyroid Res*. 2015;2015:438235.
55. de Bekker-Grob E, Bliemer M, Donkers B, Essink-Bot M-L, Korfage I, Roobol M, et al. Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. *Br J Cancer*. 2013;109(3):633–40.
56. Damen TH, de Bekker-Grob EW, Mureau MA, Menke-Pluijmers MB, Seynaeve C, Hofer SO, et al. Patients' preferences for breast reconstruction: a discrete choice experiment. *J Plast Reconstr Aesthet Surg*. 2011;64(1):75–83.
57. Ngorsuraches S, Thongkeaw K. Patients' preferences and willingness-to-pay for postmenopausal hormone receptor-positive, HER2-negative advanced breast cancer treatments after failure of standard treatments. *Springerplus*. 2015;4:674.
58. Bessen T, Chen G, Street J, Elliott J, Karnon J, Keefe D, et al. What sort of follow-up services would Australian breast cancer survivors prefer if we could no longer offer long-term specialist-based care? A discrete choice experiment. *Br J Cancer*. 2014;110(4):859–67.