A Systematic Review to Identify the Use of Preference Elicitation Methods in Healthcare Decision Making

Marieke G. M. Weernink · Sarah I. M. Janus ·
Janine A. van Til · Dennis W. Raisch ·
Jeannette G. van Manen · Maarten J. IJzerman

Abstract

Introduction   Preference elicitation methods help to increase patient-centred medical decision making (MDM) by measuring benefit and value. Preferences can be applied in decisions regarding reimbursement, including health technology assessment (HTA); market access, including benefit–risk assessment (BRA); and clinical care. These three decision contexts have different requirements for use and elicitation of preferences.

Objectives   This systematic review identified studies using preference elicitation methods and summarized methodological and practical characteristics within the requirements of the three contexts.

Methods   The search terms included those related to MDM and patient preferences. Only articles with original data from quantitative preference elicitation methods were included.

Results   The selected articles ($n = 322$) included 379 preference elicitation methods, comprising matching methods [MM] ($n = 71$, 18.7 %), discrete choice experiments [DCEs] ($n = 96$, 25.3 %), multi-criteria decision analysis ($n = 12$, 3.2 %) and other methods ($n = 200$, 52.8 %; i.e. rating scales, which provide estimates inconsistent with utility theory). Most publications of preference elicitation methods had an intended use in clinical decisions ($n = 134$, 40 %). Fewer preference studies had an intended use in HTA ($n = 68$, 20 %) or BRA ($n = 12$, 4 %). In clinical decisions, rating, ranking, visual analogue scales and direct choice are used most often. In HTA, DCEs and MM are both used frequently, and elicitation of preferences in BRA was limited to DCEs.

Conclusion   Relatively simple preference methods are often adequate in clinical decisions because they are easy to administer and have a low cognitive burden. MM and DCE fulfil the requirements of HTA and BRA but are complex for respondents. No preference elicitation methods with a low cognitive burden could adequately inform HTA and BRA decisions.

Marieke G. M. Weernink and Sarah I. M. Janus contributed equally to this work.

M. G. M. Weernink (✉) · S. I. M. Janus ·
J. A. van Til · J. G. van Manen · M. J. IJzerman
Health Technology and Services Research, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: m.g.m.weernink@utwente.nl

D. W. Raisch
College of Pharmacy, University of New Mexico, MSC 09 5360, Albuquerque, NM 87131, USA

Key Points

Preference elicitation methods can be used to quantify relative benefits and to value various aspects of a drug or health states.

Most studies found in the current literature identify preferences for guiding clinical decisions.

Fewer preference studies directly support reimbursement (health technology assessment [HTA]) or market access (benefit–risk assessment [BRA]) decisions.

Clinical decisions require more patient-friendly and straightforward preference methods.

Matching methods and discrete choice experiments fulfil almost all of the contexts' requirements of BRA and HTA. However, those methods can be cognitively complex for the respondents if the number of attributes is large or the attributes are difficult to understand.

1 Introduction

In recent years, patients, policy makers and health professionals have expressed a desire for greater patient and public involvement in healthcare decision making [1, 2]. Patient involvement in clinical decision making (CDM) has been widely encouraged [3]; it is a collaborative process between the physician and the patient. Patient involvement in health technology assessment (HTA) and benefit–risk assessment (BRA) is thought to ensure that any decisions that are made increase patient welfare by meeting patient needs and desires [4–6]. Moreover, not only patients but also the public—i.e. citizens and taxpayers—need to be engaged in policy decision making [4]. Presently, patient involvement in decision making is usually operationalized through direct involvement of the patient or a patient representative in a group of decision makers, these being either health professionals in clinical care or experts in decision panels deciding on reimbursement or drug market access at the policy level [7]. More recently, indirect involvement of the patient and the public perspective on health innovations has received increasing attention. Indirect elicitation of the patient perspective is proposed through various methods that aim to measure the usefulness of health innovations for patients (or the public) in a representative sample of the population. The results are presented to decision panels for consideration during decision making. Indirect involvement is thought to increase the representativeness of the patient perspective in deliberation on the desirability of adopting a particular health innovation [6].

The proposed methods for indirect elicitation of the patient perspective on the value of health innovations vary widely in terms of their methodological characteristics and practical applications. For instance, patient-reported outcomes (PROs) are a well-established method used to evaluate health innovations from the perspective of the patient [8]. The simplest PROs use various types of rating scales to measure the effectiveness of an innovation, especially with regard to distinct aspects of a patient's health. The health-related quality-of-life (HRQoL) measurement offers the opportunity to value health innovations in several aggregated domains of health. PROs focus on identifying the state of health or the increase in health as a result of an intervention, from the perspective of the patient. However, most PROs, including the HRQoL measurement, focus on health gains or losses resulting from the healthcare innovation but do not address the 'utility' of an innovation across domains [8]. Estimating the utility of outcomes is central to the concept of preference methods and differentiates preference methods from PROs.

Because there is a large variety of preference methods, it is often difficult to determine which concept of utility is being considered by the authors of a study. In the neoclassical utility theory, utility has ordinal properties and is regarded as a measure of the strength of a preference regarding an outcome. The concept of the expected utility theory, as defined by von Neumann and Morgenstern [9], is often used in health economics. In this theory, the measurement of utility is based on preferences regarding random health outcomes and results in a cardinal measure of utility. In response to this confusion, Carson and Louviere [10] proposed a common nomenclature for preference methods. In their nomenclature, preference elicitation methods are, first and foremost, categorized on the basis of the nature of the information that is collected, and methods are defined as (a) matching methods (MM); (b) discrete choice experiment (DCE); and (c) other methods. Although it was developed from an environmental economics perspective, this nomenclature has been previously applied in healthcare research [11]. The key difference between MM and DCE is the valuation response. In MM, respondents are asked to provide a number (or numbers) that will make them indifferent to the good outcome to be valued and the number of an explicit valuation scale. Valuation scales can be monetary, such as willingness to pay (WTP), or may include time spent in a health state, such as in time trade-off (TTO). A specific subset of MM includes standard gamble (SG), in which uncertainty, operationalized as the risk of death, is explicitly incorporated into the valuation task. DCEs belong to another category and are considered as "a general preference elicitation approach that asks respondents to make choice(s) between two or more discrete alternatives where at least one attribute of the alternative is systematically varied across respondents in such a way that information related to preference parameters of an indirect utility function can be inferred" [10]. The DCE measurement scale is an implicit cardinal utility scale. Various DCE methods can be distinguished on the basis of the format of the question. For instance, respondents can be asked to provide a binary choice for an alternative (binary choice experiment [BCE]) or a multinomial choice between various scenarios (multinomial choice experiment [MCE]), or they can provide a complete ranking of a set of alternatives based on preferences. In best–worst choice experiments (BWS), respondents are asked to identify both the best and worst alternatives from a set. In addition to the different preference methods identified by Carson and Louviere [10], multi-criteria decision methods (MCDM) are increasingly being used to explicitly consider patient and/or public preferences in decision making in healthcare. While preferences regarding the characteristics of the innovation are

calculated decompositionally in most preference elicitation methods, i.e. the respondent values the innovation as a whole, and the impact of the various characteristics is calculated from these overall valuations, MCDM offers a compositional approach to preference elicitation. In the compositional methods, respondents evaluate (pairs of) criteria separately, after which an overall composite value for the innovation is estimated [12, 13].

Generally, MM-elicited utilities have cardinal properties, while preferences resulting from DCEs have ordinal properties. Other methods, such as MCDM and those discussed in the article published by Carson and Louviere [10] (e.g. visual analogue scales [VAS] and numerical rating scales), do not provide utilities in the economic sense, because the concept of risk is not incorporated. These methods were developed on pragmatic grounds rather than being underpinned by theoretical assumptions.

Preference methods offer the potential to increase patient-centred healthcare decision making by offering some measure of benefit along with some measure of value. Several reviews of preference methods have already been published [14–17]. The studies by Hauber et al. [15, 16] and Dolan [15, 16] also address the context of use. Our paper follows those two reviews but also provides an extensive overview of methods and their use in healthcare decision making. The objective of this paper is to identify studies that used a preference elicitation method, representing the patient's view. This includes methods that (a) measure multiple domains of health or healthcare; (b) require trading between desirable and undesirable aspects of health innovation (also called attributes); and (c) offer a formal approach to valuing the underlying treatment components. The second objective is to identify the intended use of the results of the studies for three types of healthcare decisions: (a) clinical decisions; (b) HTA/reimbursement decisions; and (c) BRA/market access decisions. Finally, the suggested fit of the preference methods actually used in decision-making contexts is discussed according to the demands of the decision context.

2 Methods

2.1 Search Strategy and Selection of Articles

A literature search was performed in Scopus and Web of Science in January and February 2014. The basic search string included the following keywords: 'patient preference', 'patient value', 'patient perspective' or 'patient choice' in the paper titles. Several database features, such as truncation, proximity operators and phrase searching, were used (if available in the search engine) to increase the relevancy of retrieval of the free-text search. This basic string was combined with another search string, identifying which healthcare decision was mentioned in the title, abstract or keywords: ('benefit risk' OR 'risk benefit'), ('health technology assessment' OR 'reimbursement'), ('shared decision making' OR 'clinical decision making') or ('medical decision making' OR 'healthcare decision making'). Publications had to be written in English and had to be labelled as original research articles. The abstracts were read by two authors independently (M.W. and S.J.) so as to minimize selection and data bias. Disagreements were resolved during a discussion between the co-authors. Articles were included if they used a quantitative preference method to elicit patient and/or public preferences regarding healthcare innovations and linked the use of such a method to support decision making. Conversely, articles were excluded if they had not collected preference data or used qualitative methods to collect preferences. In cases of uncertainty of categorization, the entire articles were read. The search strategies are described in detail in Electronic Supplement 1.

2.2 Data Extraction and Interpretation

The data extraction and interpretation consisted of four steps. In step 1, all of the methods identified in the articles were described on the basis of their methodological characteristics. In step 2, the articles were categorized on the basis of the intended practical use of the results in the preference study. In step 3, all of the methods were judged on three criteria related to practical application, to identify the strengths and weaknesses of each method. Finally, in step 4, the practical use of the methods in the various decision contexts (step 2) was matched with the demands of each decision context, in terms of the identified methodological demands (step 1) and the practical demands (step 3). A more detailed description of the steps is provided below.

In step 1, the methods that were found were categorized according to the type of preference method distinguished by the nomenclature provided by Louviere and Carson [10] (MM, DCE, MCDM or other). The methods were ascribed correspondingly with the following methodological properties: direct or indirect weighting of criteria, using single or sequences of questions and outcome measures. Additionally, the need for explicit trading behaviour on the part of the respondent was described. Trading is thought to imitate the real-life context, where no result is obtained without costs, whether monetary costs or potential health loss. Explicit trade-offs can be made by trading outcomes with various characteristics of the innovation, such as a benefit–risk trade-off, or by reference to good versus bad trade-offs to gain health, i.e. the probability of unfavourable outcomes, such as the risk of death or a reduced

lifespan, as in TTO. A methodological categorization of preference methods was provided by the use of key references.

In step 2, we distinguished intended from actual use in CDM, HTA and BRA. In CDM, preferences are used to support decision making by offering a measure for patients' values. In HTA, public or patient preferences are used to support reimbursement decisions, and in BRA, patients' values are elicited concerning the perceived benefits and risks of a health innovation, thus advising market authorization. If no use was described by the original authors or the use could not be deduced by the reviewers, the article was classified as 'decision context not specified'. Subsequently, the attributes used to describe healthcare innovation were extracted. According to Ryan et al. [18], the attributes of health innovations can be classified as either health outcomes, non-health outcomes or process characteristics. Health outcomes describe the impact of an innovation on patients' health in terms of factors such as effectiveness, the risks involved, symptom relief and life expectancy. Non-health outcomes entail other attributes of care that are related to patient benefit, such as information provision and attention given to the patient and their carers. Process characteristics describe the physical characteristics of the innovation, such as the mode of drug administration, waiting time, location, frequency of consultations and costs to the patient or society.

In step 3, three criteria were used to judge the practical application of the methods, identifying their strengths and weaknesses. These criteria were 'cognitive effort on the part of the respondent', 'costs of data collection' and 'skills required in data analysis and interpretation'. 'Cognitive effort' assesses the perceived difficulty of the method for respondents. Methods that require single choices or few rankings or ratings were categorized as requiring low cognitive effort, while methods that require trading between multiple criteria were categorized as requiring high cognitive effort. The costs of data collection entail all practical demands regarding data collection. The need for large sample sizes to calculate reliable estimates, collecting data by interviews and large time investments are seen as practical difficulties. Small sample sizes and administration via online surveys are considered as presenting low or medium practical difficulties, depending on the estimated time investments.

The 'skills required in data analysis and interpretation' describe the level of difficulty in applying and/or understanding a method and are categorized as (a) basic (do not require medical/statistical expertise and do not require specialist software to implement); (b) intermediate (do not require extensive medical/statistical expertise but may require implementation of specialist software); or (c) advanced (require extensive medical/statistical expertise and may require implementation of specialist software).

Wherever possible, information about these three criteria was extracted from articles comparing specific preference methods. Additional scientific literature regarding comparison of these specific methods was consulted when the reference list was considered to be insufficient. Ultimately, if there was no literature on a particular method, a judgment of the method on the criteria was determined on the basis of agreement between all authors of this paper.

In step 4, the practical use of the methods in the three decision contexts was matched with the requirements of the decision context, in terms of the methodological and practical demands identified in earlier steps. The following hypotheses were formulated for each decision context. In the CDM setting, a simple, hands-on approach is needed, which can be used for elicitation of the preference of a single patient or a group of patients. The methods should allow for inclusion of various types of attributes deemed relevant to patients, and should therefore allow for inclusion of health, non-health and process characteristics. In order to be used in clinical practice, the methods should preferably be low in cognitive effort on the part of the respondent, should be low in costs and should require only basic technical skills in analysis and interpretation. In the clinical context, there is no need for formal calculation of (part-worth) utility or use of the outcomes in other benefit measures, as preferences regarding alternatives (I prefer treatment A to treatment B) will generally suffice in supporting decisions. Compared with clinical decisions, BRA and HTA decisions have more impact on society. These high-impact decisions suggest a preference elicitation method that assists in making the decision-making process transparent, legitimate and accountable. In the policy context, it is important to develop a benefit function; thus, adherence to utility theories is essential. Utilities are preferences measured under conditions of uncertainty, which gives the elicited preferences a solid methodological basis. Furthermore, HTA refers to the evidence of clinical effectiveness, safety and cost-effectiveness across groups of patients; therefore, the methods should account for several patient-relevant attributes and outcome measures. In the BRA context, drugs are approved on the basis of their safety and effectiveness; thus, the method needs to incorporate multiple health outcomes. For both BRA and HTA decision-making contexts, the preference elicitation methods should be easy and simple for patients to understand. In comparison with CDM, there is more time for collecting preferences, because the approval process takes months or even years. Furthermore, there are more resources and time for collection and analysis of preferences. Methods with more advanced statistics are therefore possible.

3 Results

3.1 Results of Search Strategy

The search strategy identified 1,036 unique articles (see Fig. 1). The initial selection revealed 322 articles that used one or more quantitative methods to elicit patients' preferences; these are listed in Electronic Supplement 2.

3.2 Patient Preference Methods

The 322 articles selected for this review were concerned with 379 preference methods (in some articles, multiple methods were described). In accordance with the classification provided by Carson and Louviere [10], 71 of these preference methods were identified as MM, 96 as DCEs, 200 as 'other methods' and 12 as MCDM. The four main categories of preference methods could be further subdivided into 15 conceptually different preference methods (Fig. 2).

3.2.1 Matching Methods

Three conceptually different methods within MM were found in the literature. This review identified 47 studies where respondents were asked with a direct question (DQ) to state their WTP to obtain a certain health state. For example, Esfandiari et al. [19] assessed the effect of long-term financing on preferences regarding implant overdentures. Seven articles used the specific WTP method. SG, another variant of DQ, was found 34 times and is used mainly to quantify patient preferences regarding health states, not only for the probability of experiencing those health states but also for decision analysis, as shown by a

study by Montgomery et al. [20]. DQ is also operationalized using a binary response, instead of a matching variable (time, money), as shown in the study by Gyrd-Hansen and Kristiansen [21], where respondents were presented with hypothetical therapy scenarios that involved life-year gains and were asked to state whether they should follow the therapy (yes/no). Another category in MM is the TTO approach, which was identified 19 times in this review. As with SG, TTO is also used to quantify the utilization of patient preferences regarding health states in decision analysis [22]. The final MM category is the allocation game, in which respondents are asked to allocate a fixed quantity to different categories. In a study by Ubel et al. [23], respondents were asked how to allocate funds to benefit people with varying levels of disability. Within MM, single-question and sequential-question versions are possible. In most studies, the sequential versions were used in order to assess a value for various goods, i.e. the health states [24]. Only three studies were identified using a single question, valuing one single health state. MM uses a direct weighting strategy; individuals are asked to directly provide their utilities for a health state. In MM, trade-offs can be made by trading health gains against several negative outcomes, such as the risk of immediate death (SG), a reduced lifespan (TTO), a monetary equivalent (WTP) or allocation trade-offs (allocation game). MM is consistent with the expected utility theory, in that outcomes have cardinal properties and result in utilities.

3.2.2 Discrete Choice Experiments

Regarding DCE, four conceptually different methods were found in the literature: BCE, MCE, full ranking exercise and BWS. In this review, the majority (90 studies) used a BCE in their DCE methods. MCE was used only two times. It should be noted that some BCEs included an additional opt-out option, in order to allow respondents to be nondemanders [25]. In the BCE category, 19 methods used the variant where one attribute of one of the treatment options is systematically varied until respondents switch their treatment preference to the alternative choice. This variant is used to determine treatment thresholds or switching points as, for example, in a study by Brundage et al. [26], where the treatment threshold represented the minimum survival percentage required by the participant in order to go for the more toxic treatment. One article in our review used a full ranking exercise. According to a study by Singh et al. [27], a full ranking approach is more realistic (i.e. it includes all hypothetical scenarios/products as in a real market); however, the disadvantage is that it yields only ordinal data. A less extensive ranking exercise is BWS, which can be used in several forms. Our review found two articles that used the BWS 'profile case', in which
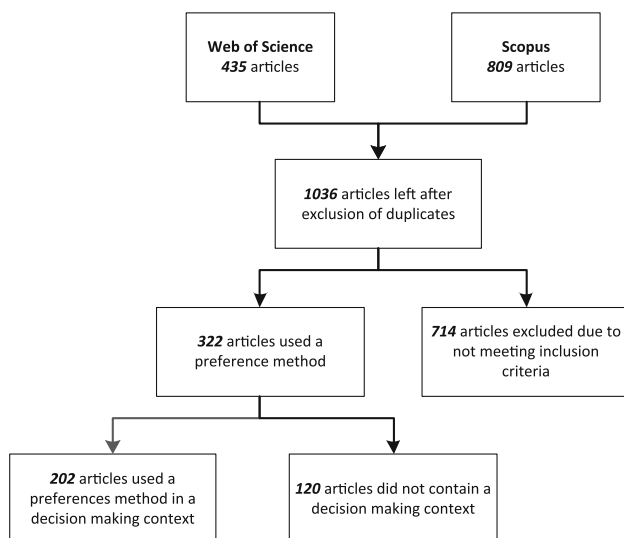

Fig. 1 Results of search strategy

Web of Science
435 articles

Scopus
809 articles

1036 articles left after exclusion of duplicates

322 articles used a preference method

714 articles excluded due to not meeting inclusion criteria

202 articles used a preferences method in a decision making context

120 articles did not contain a decision making context

Fig. 2 Patient preference methods found in the literature. * Within the 'direct questions' category, seven articles used willingness to pay and 34 articles used a standard gamble
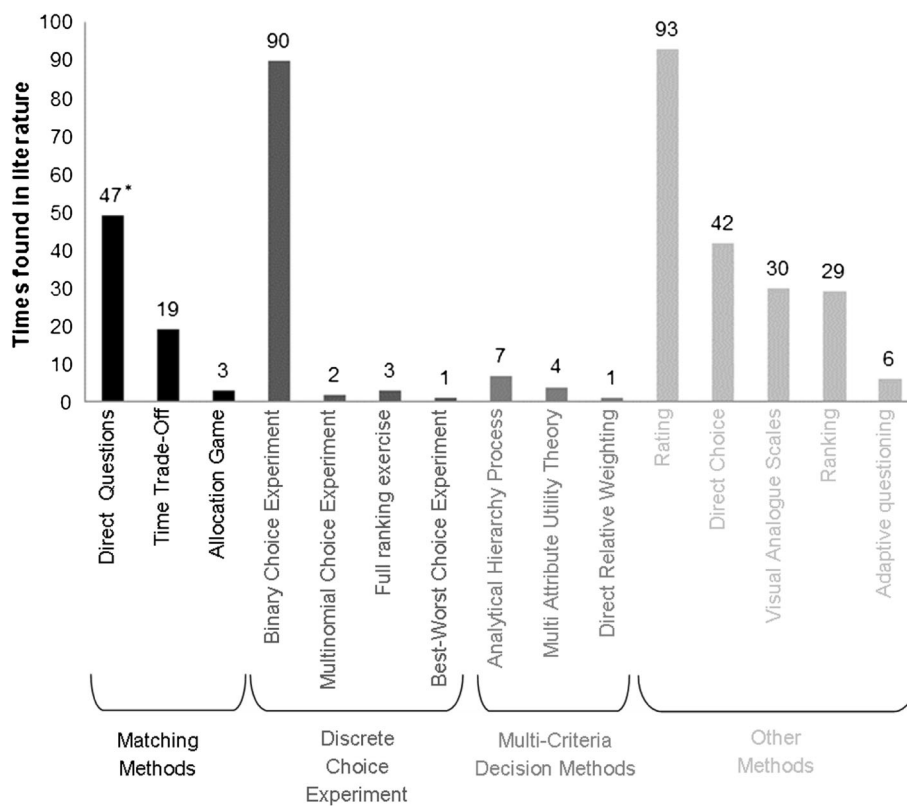
respondents judged each scenario at a time and selected the most attractive and the least attractive attribute. Also, one 'multi-profile case' was found where preferences were elicited by selecting the most preferred and least preferred scenarios out of a set of three or more alternatives [28, 29]. In contrast to MM, DCEs are mostly used to estimate the marginal value of changing attributes of the health state to be judged through sequential questioning. No preferences regarding single stand-alone scenarios were found by our review. A unique aspect of DCE is that individuals are asked to provide their utilities for each attribute not directly but indirectly by determining relative preferences regarding choice scenarios. Sequential questioning and subsequent regression-based analysis give overall utility and individual part-worth utilities or an estimate of the marginal value of changing attributes.

3.2.3 Multi-Criteria Decision Methods

Three different MCDMs were found in this review: the analytic hierarchy process (AHP), multi-attribute utility theory (MAUT) and direct relative weighting (DRW). All MCDMs are compositional but differ in other characteristics. MAUT, an extension of the expected utility theory of multiple criteria, calculates the expected utility of an innovation by mapping criterion-specific measurement scales onto a common scale of utility [30]. The order of the

importance of attributes is determined by presenting explicit trade-offs between gambles to the decision maker, resulting in direct assignment of weights to the decision criteria (overall utility and part-worth utility weights). Only four studies used MAUT, e.g. to model prostate cancer patients' preferences regarding health states [31]. AHP does not assign weights directly but presents respondents with pairwise comparisons between criteria. A value scale that preserves the ratio information contained in the pairwise comparisons can then be estimated. Hummel et al. [32] used AHP to elicit patient preferences regarding multiple outcome measures to prioritize antidepressant drug treatment. This review identified seven studies using AHP. In DRW, respondents are asked to assign numbers to every decision characteristic being assessed. Summary scores (typically created using the weighted average method) of direct weighting indicate how well the alternative meets the goal. DRW does not exhibit explicit trading of characteristics, and only criteria weights are calculated. One study employed this preference method [33].

3.2.4 Other Methods

According to Carson and Louviere [10], this category covers (amongst others) methods that cannot give estimates consistent with utility theory. Furthermore, it includes

methods that use self-explicated ratings. Those methods are problematic because the resulting data are ordinal but are sometimes used as interval. Methods such as adaptive questioning, rating, ranking, VAS and direct choice fall into this category. Rating, ranking and VAS are compositional methods and use a single question to directly elicit preferences regarding health states [34]. A VAS is a sliding scale with anchored end points. In the review, 30 studies used a VAS to value health scenarios. They utilized 1–100 [35] or two statements (no information, full information) as anchored end points [36]. Numerical rating scales are used to value health states on a scale ranging from not important (0) to very important (10) [37] or 0–4 [38]. Rating scales were used in 93 studies. Ranking exercises (also known as comparative scales) are used to make relative judgments of preference by asking respondents to indirectly weight (rank) health states and put them in their preferred order based on their value. In a study by Guerlain et al. [39], participants were asked to rank the four products by moving them on the table into the preferred ranking order from left to right. The ranking method was used in 29 studies. Adaptive questioning is a combination of self-explicated ratings and BCE, and each question is based on the respondent's answer to previous questions. The indirect weighting strategy of adaptive questioning results in part-worth weights for each attribute. This review identified six studies using adaptive questioning. Direct choice is a binary choice between two health states to determine which is valued highest. It is compositional, and the weights are drawn up directly. Our review identified 42 studies using direct choice questions in one form or another, ranging from questionnaires consisting only of direct choice questions [40] to studies using direct choice as an additional method—for example, with a VAS [41].

3.3 Preference Elicitation Use in Healthcare Contexts

The 322 identified articles were classified according to their intended use in three healthcare decision-making contexts. Four percent of the studies were classified as BRA ($n = 12$), 20 % were classified as HTA ($n = 68$), 40 % were classified as CDM ($n = 134$) and 36 % of the articles specified no intended context of use ($n = 120$). The intended use was not specified among the methods evenly. Eleven studies addressed two contexts of decision making, and these were counted twice. We classified one article as indicating both BRA and HTA, and ten articles indicated both CDM and HTA decision-making contexts (Table 1).

3.3.1 Benefit–Risk Assessment

All of the articles within BRA used DCEs to elicit preferences (Table 1). Patients' maximum acceptable risk (MAR) for each unit of benefit [42, 43] or the minimal accepted benefit (MAB) [44], were calculated from the preference data obtained by the DCEs. MAR and MAB provide decision makers with information about the payoff that patients require to accept risk or which minimal clinical benefit is required from a healthcare innovation to make it acceptable to patients. Studies designed for or aimed at the BRA context elicit preferences regarding health outcomes (Table 2). Our literature review showed that process characteristics and non-health outcomes are often not taken into account in BRA.

3.3.2 Health Technology Assessment

In the reimbursement context, there is no clear preference for any one method (Table 1). Patient preferences in economic analysis are mainly measured with MM. Preferences measured with TTO and SG are often used in economic evaluations within the cost–utility framework; quality-adjusted life-years (QALYs) are used in decision analysis. Examples of studies with economic evaluations of preference methods are those by Sanelli et al. [45–47], Carroll and Downs [45–47] and Netten et al. [45–47]. Understanding of patient and societal preferences and resource allocation decisions has gained importance because reimbursement decisions are increasingly based on economic evaluations [48]. Our results show that various preference

Table 1 Intended use of preference methods in decision-making contexts

	Benefit–risk assessment	Health technology assessment	Clinical decision making	Decision-making context not specified
Matching methods	–	19	23	32
Discrete choice experiment	12	24	31	30
Multi-criteria decision methods	–	3	3	7
Other methods[a]	–	36	98	74
Total number of articles in context	12 (1)[b]	68 (11)[b]	134 (10)[b]	120

[a] Includes visual analogue scales, direct choice, adaptive questioning, numerical rating and ranking

[b] The numbers in parentheses indicate the numbers of articles that imply an intended use for two contexts and thus the methods were counted twice

Table 2 Attribute valuation in decision-making contexts

	Health outcomes	Non-health outcomes	Process characteristics	Total number of articles in context
Benefit–risk assessment	12	–	3	12
Health technology assessment	51	28	41	68
Clinical decision making	103	48	44	134

methods for priority setting of resource allocation have been used. The studies have often applied a resource-allocation trade-off, which can assist the decision maker in making a choice between competitive fundings of health innovations. For example, Mason et al. [49, 50] used BCE to elicit public preferences regarding prioritization of healthcare interventions in the UK, and Ratcliffe [49, 50] employed rankings to elicit preferences from the recipients of donor liver grafts. Furthermore, patient preference methods are also used to compare patient preferences regarding competing therapies, e.g. two comparing chemotherapies for breast cancer [51] or two insulin types for diabetes [52]. Preferences regarding certain process characteristics, such as (out-of-pocket) costs and mode of administration, are elicited more often in HTA than in BRA (Table 2). Process characteristics can have an impact on adherence to a therapy; various articles in this review state that because of non-adherence, potential clinical benefits are lost and that the costs rise per QALY, which is a relevant outcome for policy decision makers [53–55].

3.3.3 The Clinical Decision Context

Fifty percent of CDM articles had rating, ranking or VAS scales to measure patient preferences (Table 1). These scales are used to measure patient preferences to support decision making in the physician's office. For instance, in a study by Witticke et al. [56], rating scales were used by physicians to elicit patient preferences to specify medication-related factors of interest. Rating scales are also utilized as a value elicitation tool in decision aids if and when multiple outcomes of treatment are taken into account, as was done in a study by Fiks et al. [38]. Methods such as SG and TTO have rarely been used in the physician's office [22, 57]. Some articles have addressed the discrepancies in preferences between physicians and patients [58]. The objective of these studies was to identify any discrepancies and to raise awareness among physicians about the relevance of preferences in decision making. Non-health outcomes and process characteristics are very important in the CDM context, as can be seen from the frequency with which they have been included in the preference instrument (Table 2). For instance, mode of administration and (out-of-pocket) costs directly influence decisions made by the

patient, which can influence uptake or compliance [56, 59]. Preferences regarding non-health outcomes, such as information and participation, are important because of their influence on the patient–physician relationship and how the treatment is delivered to the patient [36, 60, 61].

3.4 Comparison of Context Requirements and Method Characteristics

In this section, the requirements of each context as outlined in the 'Methods' section are compared with the methodological and practical characteristics of the methods currently used in the three decision-making contexts. Electronic Supplement 3 shows the scores of each preference method for previously defined methodological and practical characteristics.

The requirements of preference elicitation and their effect on CDM show a good fit with the actual use of preference methods. The choice of relatively simple methods (such as ranking, rating, VAS and direct choice) in the clinical decision context is explained by their ease of use, the low cognitive burden on the patient and the clinician, and the ease of estimating preferences from the value exercise. Ease of administration and fast access to results allow physicians or other healthcare professionals to utilize the results of the preference exercise in clinical practice. In BRA/HTA decisions, MM and DCEs are used the most, and these methods fulfil the theoretical and data collection and analysis requirements in these contexts. MM and DCEs assist in making the decision-making process transparent, legitimate and accountable. Both MM and DCEs have strong methodological underpinnings in utility and economic theory. However, DCEs can be cognitively complex for respondents if the number of attributes is large or the attributes are difficult to interpret (probabilities). Similarity, SG and TTO are cognitively complex because they deal with abstractions of life and death, and they vary the life-years and health conditions by eliciting preferences. The compositional approach of MCDMs might be easier because preferences are elicited for each attribute at a time. However, MCDM does not provide utilities. MAUT (the only MCDM found by our review that provides utilities) is judged as being exhausting for the respondent because it demands a vast amount of preference information from patients [62].

4 Discussion

The aim of this systematic review was to identify studies that used a preference elicitation method and to identify the healthcare decision context of the studies. Additionally, the context requirements for eliciting preferences in three separate decision contexts were identified and compared with the actual use of preferences.

In total, 379 methods were identified in 322 articles representing 15 preference elicitation methods. Four percent of the studies were classified as BRA, 20 % were classified as HTA, 40 % were classified as CDM and 36 % of the articles did not specify an intended context of use. Regarding clinical decisions, most preferences are elicited by means of straightforward methods such as rating, ranking, VAS and direct choice. Our findings show a good fit between the actual use of methods and the hypothesized demands for preferences in clinical decisions. Although preferences elicited with relatively simple methods do not result in exact utilities, utilities are not required in clinical decisions, where the results mainly need to inform on relative treatment preferences (e.g. for treatment A versus treatment B). In contrast, the methods used in BRA and HTA need to adhere to the leading paradigm of utilities in health economics [63] to ensure strong methodological underpinnings. For example, DCEs are less prone to bias than rating or VAS [64]. MM and DCEs fulfil the theoretical requirements needed in HTA and BRA. However, both MM and DCE can be cognitively complex for the respondents, and this has been noted as a limitation in multiple studies [43, 65, 66].

Although MM is a recognized and long-standing method for preference elicitation, few MM methods were identified in this review. MM is generally used in QALY measurement for performing cost-effectiveness analysis from an HTA perspective, not explicitly to incorporate individual patient values. MCDMs (AHP, MAUT) were the least reported preference methods, even though they can also be used to accommodate the weighting process of multiple conflicting objectives in the HTA context. MCDM could be used to weigh the importance of the various sources of evidence in the decision process. Hence, employing MCDM can result in greater transparency of the decision process and increased patient centredness. However, there is debate as to whether MCDM can be used as a true preference method or whether it would be better used as a decision-support tool [67, 68].

There are some limitations of this review, which influence the interpretation of its outcomes. First, there is a large variety of preference methods in the literature, and several categorizations of those preference methods exist. In this review, we chose to use the nomenclature of Carson and Louviere [10] to frame the preference methods. However, this is not a universally accepted nomenclature.

Second, the methods were characterized according to their cognitive and practical difficulties. However, no publications that compared different methods on the basis of those characteristics were found. In any case, both characteristics depend on how they are used in specific studies. The cognitive complexity of BCE and MCE, for example, is significantly affected by an increase in the number of attributes [69]. Therefore, a clear categorization depends on one's own evaluation and opinion.

Third, as with any systematic review, the search terms and selection process influenced which studies were included. Earlier reviews differed with regard to their search terms or explicitly used the name of the preference method in the search terms. The current review found that only DCEs were used to inform BRA, but Hauber et al. [15] stated that benefit–risk preferences can also be measured through direct elicitation methods (rating scale, threshold technique and SG). However, in agreement with the results of this article, they also stated that DCEs are better tailored to estimate benefit–risk preference weights for multiple outcomes simultaneously.

5 Conclusion

We identified studies that actually only used their results in practice in the CDM context—for example, for development of a decision aid [33, 70, 71]. In most articles, explicit or implicit statements made by the authors referred to just an intended use of the results. In the current state of the art, there is mainly a methodology push; no empirical studies of use of preferences to support decisions in BRA/HTA have been published. The importance of the patient's view when making clinical decisions was acknowledged a long time ago. Healthcare is increasingly concerned with the desires of patients, and patients are encouraged to participate in decision making. This awareness of patient centredness facilitates the use of patient preferences in the clinical context. Preferences regarding health outcomes, and thus the value of outcomes to the patient and society, are also of increasing importance in other decision contexts. Currently, BRA and HTA agencies have established procedures for evaluating (new) health innovations. Inclusion of patient preferences would provide an extra source of evidence for the existing value dossier required for drug approval or reimbursement. Yet, this would require agencies to adapt their procedures to determine how patient preferences can best be measured and how they will fit into this process. Therefore, it is necessary to establish which method best fits the context demands. However, this review indicates a tension between the need

to underpin methodological preferences and the cognitive and practical difficulties that come with those preference methods. No method exists that has a low cognitive burden and strong methodological underpinnings and that can, at the same time, deliver enough and adequate information to support the decision.

References

1. Litva A, et al. 'The public is too subjective': public involvement at different levels of health-care decision making. Soc Sci Med. 2002;54(12):1825–37.

2. Coulter A. Patient engagement—what works? J Ambul Care Manag. 2012;35(2):80–9.

3. Smith M, Higgs J, Ellis E. Factors influencing clinical decision making, in clinical reasoning in the health professions. Oxford: Butterworth-Heinemann; 2008. p. 89–100.

4. Bridges J, Jones C. Patient-based health technology assessment: a vision of the future. Int J Technol Assess Health Care. 2007;23(1):30–5.

5. Gagnon M, et al. Introducing patients' and the public's perspectives to health technology assessment: a systematic review of international experiences. Int J Technol Assess Health Care. 2011;27(1):31–42.

6. van Til J, Ijzerman M. Why should regulators consider using patient preferences in benefit–risk assessment? Pharmacoeconomics. 2014;32(1):1–4.

7. European Medical Agency (2013). The patient's voice in the evaluation of medicines: how patients can contribute to assessment of benefit and risk. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2013/10/WC500153276.pdf. Accessed 11 Jul 2013.

8. Bridges J, et al. Patient preference methods—a patient centered evaluation paradigm. ISPOR Connect. 2007;12(6):4–7.

9. von Neumann J, Morgenstern O. Theory of games and economic behavior. Princeton: Princeton University Press; 1944.

10. Carson R, Louviere J. A common nomenclature for stated preference elicitation approaches. Environ Resour Econ. 2011;49(4):539–59.

11. Ali S, Ronaldson S. Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. Br Med Bull. 2012;103:21–44.

12. Meißner M, Decker R. An empirical comparison of CBC and AHP for measuring consumer preferences. In: International Symposium of Analytical Hierarchy Process; 2009; Pittsburgh.

13. Green P, Srinivasan V. Conjoint analysis in consumer research: issues and outlook. J Consum Res. 1978;5:103–23.

14. Louviere J. Conjoint analysis modelling of stated preferences: a review of theory, methods, recent developments and external validity. Journal of Transport Economics and Policy. 1988;22(1):93–119.

15. Hauber A, Fairchild A, Johnson F. Quantifying benefit–risk preferences for medical interventions: an overview of a growing empirical literature. Appl Health Econ Health Policy. 2013;11(4):319–29.

16. Dolan J. Multi-criteria clinical decision support. Patient. 2010;3(4):229–48.

17. Ryan M, et al. Eliciting public preferences for healthcare: a systematic review of techniques. Health Technol Assess. 2001;5(5):186.

18. Ryan M, Skåtun D, Major K. Using discrete choice experiments to go beyond clinical outcomes when evaluating clinical practice. In: Ryan M, Gerard K, Amaya-Amaya M (eds). Using discrete choice experiments to value health and health care,. Dordrecht: Springer; 2008:101–16.

19. Esfandiari S, et al. Implant overdentures for edentulous elders: study of patient preference. Gerodontology. 2009;26(1):3–10.

20. Montgomery AA, Harding J, Fahey T. Shared decision making in hypertension: the impact of patient preferences on treatment choice. Fam Pract. 2001;18(3):309–13.

21. Gyrd-Hansen D, Kristiansen IS. Preferences for 'life-saving' programmes: small for all or gambling for the prize? Health Econ. 2008;17(6):709–20.

22. Sommers B, et al. Decision analysis using individual patient preferences to determine optimal treatment for localized prostate cancer. Cancer. 2007;110(10):2210–7.

23. Ubel P, et al. Public preferences for prevention versus cure. Med Decis Mak. 1998;18:141–8.

24. Hallan S, et al. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. J Intern Med. 1999;246(3):309–16.

25. Marti J. Assessing preferences for improved smoking cessation medications: a discrete choice experiment. Eur J Health Econ. 2012;13(5):533–48.

26. Brundage MD, et al. Using a treatment-tradeoff method to elicit preferences for the treatment of locally advanced non-small-cell lung cancer. Med Decis Mak. 1998;18(3):256–67.

27. Singh J, et al. Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. Med Care. 1998;36(8 Suppl):AS31–45.

28. Swancutt DR, Greenfield SM, Wilson S, Women's colposcopy experience and preferences: a mixed methods study. BMC Womens Health. 2008;8:2.

29. Al-Janabi H, Flynn TN, Coast J. Estimation of a preference-based carer experience scale. Med Decis Mak. 2011;31(3):458–68.

30. Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value trade-offs. Cambridge: Cambridge University Press; 1993.

31. Chapman GB, et al. A multi-attribute model of prostate cancer patients' preferences for health states. Qual Life Res. 1999;8(3):171–80.

32. Hummel MJM, et al. Using the analytic hierarchy process to elicit patient preferences: prioritizing multiple outcome measures of antidepressant drug treatment. Patient. 2012;5(4):225–37.

33. Cunich M, et al. Integrating evidence and individual preferences using a web-based multi-criteria decision analytic tool: an application to prostate cancer screening. Patient. 2011;4(3):153–62.

34. McCormack HM, Horne DJ, Sheather S. Clinical applications of visual analogue scales: a critical review. Psychol Med. 1988;18(4):1007–19.

35. Heckerling PS, Verp MS, Albert N. Patient or physician preferences for decision analysis: the prenatal genetic testing decision. Med Decis Mak. 1999;19(1):66–77.

36. Davis M, Hoffman J, Hsu J. Impact of patient acuity on preference for information and autonomy in decision making. Acad Emerg Med. 1999;6(8):781–5.

37. Ruland C. Decision support for patient preference-based care planning effects on nursing care and patient outcomes. J Am Med Inform Assoc. 1999;6(4):304–12.

38. Fiks A, et al. Development of an instrument to measure parents' preferences and goals for the treatment of attention deficit-hyperactivity disorder. Acad Pediatr. 2012;12(5):445–55.

39. Guerlain S, Hugine A, Wang L. A comparison of 4 epinephrine autoinjector delivery systems: usability and patient preference. Ann Allerg Asthma Immunol. 2010;104(2):172–7.

40. Fagerlin A, Zikmund-Fisher B, Ubel P. Cure me even if it kills me: preferences for invasive cancer treatment. Med Decis Mak. 2005;25(6):614–9.

41. De Boer A, et al. Transhiatal vs extended transthoracic resection in oesophageal carcinoma: patients' utilities and treatment preferences. Br J Cancer. 2002;86(6):851–7.

42. Johnson FR, et al. Quantifying women's stated benefit–risk trade-off preferences for IBS treatment outcomes. Value Health. 2010;13(4):418–23.

43. Bewtra M, et al. Patient preferences for surgical versus medical therapy for ulcerative colitis. Inflamm Bowel Dis. 2014;20(1): 103–14.

44. Hauber AB, et al. Patients' benefit–risk preferences for chronic idiopathic thrombocytopenic purpura therapies. Ann Pharmacother. 2010;44(3):479–88.

45. Sanelli PC, et al. Work-in-progress toward incorporating patients' preferences in practice guidelines for imaging aneurysmal subarachnoid hemorrhage. Acad Radiol. 2009;16(5):535–40.

46. Carroll AE, Downs SM. Improving decision analyses: parent preferences (utility values) for pediatric health outcomes. J Pediatr. 2009;155(1):21–5.

47. Netten A, et al. Outcomes of social care for adults: developing a preference-weighted measure. Health Technol Assess. 2012;16(16):1–166.

48. Desser AS, Olsen JA, Grepperud S. Eliciting preferences for prioritizing treatment of rare diseases: the role of opportunity costs and framing effects. PharmacoEconomics. 2013;31(11): 1051–61.

49. Mason H, Baker R, Donaldson C. Understanding public preferences for prioritizing health care interventions in England: does the type of health gain matter? J Health Serv Res Policy. 2011;16(2):81–9.

50. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. Health Econ. 2000;9(2):137–48.

51. Beusterien K, et al. Patient preferences for chemotherapies used in breast cancer. Int J Womens Health. 2012;4:279–87.

52. Aristides M, et al. Conjoint analysis of a new chemotherapy: willingness to pay and preference for the features of raltitrexed versus standard therapy in advanced colorectal cancer. PharmacoEconomics. 2002;20(11):775–84.

53. Hiligsmann M, et al. Health technology assessment in osteoporosis. Calcif Tissue Int. 2013;93(1):1–14.

54. Laba TL, et al. Patient preferences for adherence to treatment for osteoarthritis: the Medication Decisions in Osteoarthritis Study (MEDOS). BMC Musculoskelet Disord. 2013;14(160): 1471–2474.

55. Benjamin L, et al. Physicians' preferences for prescribing oral and intravenous anticancer drugs: a discrete choice experiment. Eur J Cancer. 2012;48(6):912–20.

56. Witticke, D et al. Do we prescribe what patients prefer? Pilot study to assess patient preferences for medication regimen characteristics. Patient Prefer Adher, 2012;6:679–84.

57. Rhodes E, et al. Preferences for type 2 diabetes health states among adolescents with or at risk of type 2 diabetes mellitus. Pediatr Diabetes. 2011;12(8):724–32.

58. Johnson F, et al. Are gastroenterologists less tolerant of treatment risks than patients? Benefit–risk preferences in Crohn's disease management. J Manag Care Pharm. 2010;16(8):616–28.

59. Vermeulen H, et al. Preferences of patients, doctors, and nurses regarding wound dressing characteristics: a conjoint analysis. Wound Repair Regen. 2007;15(3):302–7.

60. Pardon K, et al. Preferences of advanced lung cancer patients for patient-centred information and decision-making: a prospective multicentre study in 13 hospitals in Belgium. Patient Educ Couns. 2009;77(3):421–9.

61. Pardon K, et al. Are patients' preferences for information and participation in medical decision-making being met? Interview study with lung cancer patients. Palliative Med. 2011;25(1):62–70.

62. Troche CJ, Paltiel AD, Makuch RW. Evaluation of therapeutic strategies: a new method for balancing risk and benefit. Value Health. 2000;3(1):12–22.

63. Smith R, et al. The capability approach: an alternative evaluation paradigm for health economics? In: Jones AM (ed). The Elgar companion to health economics. 2nd ed. Cheltenham: Edward Elgar Publishing; 2012:415.

64. Torrance G, Feeny D, Furlong W. Visual analog scales do they have a role in the measurement of preferences for health states? Med Decis Mak. 2001;21(4):329–34.

65. Johnson F, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. Gastroenterology. 2007;133(3):769–79.

66. Johnson F, et al. Multiple sclerosis patients' benefit–risk preferences: serious adverse event risks versus treatment efficacy. J Neurol. 2009;256(4):554–62.

67. European Medicines Agency (2010). Benefit–risk methodology project. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/document_listing/document_listing_000314.jsp. Accessed 29 Oct 2013.

68. Diaby V, Goeree R. How to use multi-criteria decision analysis methods for reimbursement decision-making in healthcare: a step-by-step guide. Expert Rev Pharmacoecon Outcomes Res. 2014;14(1):81–99.

69. Johnson FR, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health. 2013;16(1):3–13.

70. Goel V, et al. Randomized trial of a patient decision aid for choice of surgical treatment for breast cancer. Med Decis Mak. 2001;21(1):1–6.

71. Ruland C. Decision support for patient preference-based care planning effects on nursing care and patient outcomes. J Am Med Informat Assoc. 1999;6(4):304–12.