BMJ Open

Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of Discrete Choice Experiments

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014719
Article Type:	Research
Date Submitted by the Author:	13-Oct-2016
Complete List of Authors:	Harrison, Mark; University of British Columbia, Pharmaceutical Sciences; UBC Milbers, Katherine; Centre for Health Evaluation and Outcome Sciences Hudson, Marie; Jewish General Hospital, Lady Davis Institute for Medical Research Bansback, Nick; School of Population and Public Health, University of British Columbia
Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Health services research, Research methods, Health economics
Keywords:	Discrete Choice Experiment, Health Care Preferences, Shared Decision-Making

Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of Discrete Choice Experiments

Mark Harrison[1,2], Katherine Milbers[2], Marie Hudson[3,4,5], Nick Bansback[2,6]

[1] Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC

[2] Centre for Health Evaluation and Outcome Sciences, St Paul's Hospital, Vancouver, BC

[3] Department of Medicine, McGill University, Montréal, QC

[4] Division of Rheumatology, Jewish General Hospital, Montréal, QC

[5] Lady Davis Institute for Medical Research, Montréal, QC

[6] School of Population and Public Health, University of British Columbia, Vancouver, BC

Correspondence to:

Mark Harrison,

Faculty of Pharmaceutical Sciences,

University of British Columbia,

2405 Wesbrook Mall

Vancouver,

BC Canada V6T 1Z3

+1 604-827-0687

mark.harrison@ubc.ca

Objectives: The aim of this paper was to review studies comparing patient and health care provider preferences for health care interventions using discrete-choice experiments (DCEs), specifically to quantify the extent to which they demonstrate evidence of concordance of patient and healthcare provider preferences and values for treatment, and to review the methodology of DCEs to evaluate similarities, differences and rigour of their designs.

Methods: We conducted a systematic review in Medline, EMBASE, Econlit, PsycINFO and Web of Science for studies that elicited values from patient and healthcare providers using Discrete Choice Experiments (DCEs). The inclusion criteria were: peer-reviewed; complete empiric text in English from 1995-July 31st 2015; discussing a healthcare-related topic; DCE methodology; comparing patients and health care providers.

Design: Systematic review

Results: We identified 38 papers from 15 countries, exploring 16 interventions in 26 diseases/indications. Methods to analyze results, determine concordance between patient and physician values, and explore heterogeneity varied considerably between studies. The majority of studies we reviewed concluded that there was more evidence of mixed concordance and discordance (n=28) or discordance of patient and health care provider preferences (n=12) than of concordant preferences (n=4). Concordance and discordance varied within studies according to the type of attribute being considered.

Conclusions: Discordant patient and health care provider preferences for the importance of different attributes of health care interventions is common. Concordance or discordance also varies according to which attributes are being considered, highlighting that concordance should not be considered as a binary outcome but should consider all aspects jointly. DCE studies provide an excellent opportunity for determining value concordance between patients and providers, but assessment of concordance was limited by a lack of consistency in the approaches used and a lack of consideration of heterogeneity of preferences.

Strengths and limitations of this study

- Discrete Choice Experiments (DCEs) provide an excellent opportunity to understand preferences as they can break down and value different components of treatments and identify the trade-offs people make between these components.
- By looking specifically at studies using DCE methodology we are taking a focussed view of the literature on this topic.
- We have systematically reviewed a large body of work, which has attempted to understand the similarities or differences of patient and health care provider preferences using DCEs.
- We highlight a lack of consistency within and between studies, which adds difficulty to summarizing findings, but leads to recommendations for future studies.
- We synthesise coefficients between and across studies to give a sense of differences in concordance by attribute.

Funding

This work was supported by a grant from the Canadian Rheumatology Association through the Canadian Initiative for Outcomes in Rheumatology Care (CIORA).

ICME statement

MHa and NB were responsible for substantial contributions to the conception or design of the work, the analysis and interpretation of data, and drafting and critically revising the work. MHu was responsible for substantial contributions to the conception or design of the work and critically revising the work for important intellectual content. KM was responsible for substantial contributions to the acquisition of data for the work and drafting of the manuscript. All authors gave their final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We are grateful to Julia Kaal, University of British Columbia, for her assistance in revising the final draft for publication.

Data sharing statement

This systematic review draws on published research only. No new data were collected or generated in undertaking this study.

Competing interests

Mark Harrison holds the UBC Professorship in Sustainable Health Care, which is funded by Amgen Canada, AstraZeneca Canada, Eli Lilly Canada, GlaxoSmithKline, Merck Canada, Novartis Pharmaceuticals Canada, Pfizer Canada, Boehringer Ingelheim (Canada), Hoffman-La Roche, LifeScan Canada, and Lundbeck Canada.

1. BACKGROUND

Patient-centered care has been long and widely advocated[1,2]. Respectful and responsive to the clinical state, circumstances, as well as preferences and actions of the patient[3,4], it posits clinical decisions to be guided by the values of the patient. Shared-decision making has been promoted as a way to achieve a patient-centred approach by ensuring that patients are fully informed about, and consequently receive options that reflect, their personal treatment preferences[5]. The assumptions of health care providers about their patients' preferences has been termed 'preference diagnosis', and is thought to be as important as the 'medical diagnosis'. Both types of diagnoses involve inferences based on imperfect information; however, while the medical diagnosis is based on a combination of tests, imaging and specialist opinions, it is less clear what informs preference diagnosis. Furthermore, the importance and complexity of diagnosing patient preference is often overlooked, and evidence suggests that health care providers erroneously deem themselves good at diagnosing their patients' preferences[6]. Discordance between patient and health care provider preferences with regard to treatment decisions can lead to preference misdiagnosis, also called the silent misdiagnosis[7].

Preference misdiagnosis affects patients as well as health systems as evidence suggest that patients whose care more closely matches their preferences consume less health care[8]. However, evidence supporting the assumption that patients and health care providers have different preferences is not clear cut. Studies attempting to understand how and how often patient and health care provider preferences differ have shown mixed results. For example, some studies suggest no differences between patient and health care provider in terms of rank or strength of preference[9] for features of treatments, while others show similar ranks but differences in terms of strength of preference[10], or differences in terms of both rank and strength of preferences[11,12]. A recent review assessing differences between patient and health care provider preferences in health care decision-making found that, at the aggregate level, patients and health care providers had different preferences[13]. However, the evidence was limited and assessment of the discordance of preferences remains unclear and complicated due to different methodologies (e.g., time trade-off[10], conjoint analysis[11], paired comparison) different disease contexts (e.g., pap tests[9], diabetes[12], head/neck cancer[10], cardiac risk assessment[11]), and types of decision [13]. Furthermore, the conclusions drawn by these studies sought to compare aggregate, or average, patient and health care provider preferences, which could distort the true problem of heterogeneous preferences.

Discrete choice experiments (DCEs) have become the gold standard [14] for understanding preferences due to their ability to break down and value different components of treatments and services (whether these are processes, structures or outcomes[15–17]) as well as identify the trade-offs people make between these different components[18,19]. Although DCEs ask about hypothetical choices, their agreement with actual choices is good. DCEs therefore facilitate realistic assessment of the concordance of patient and health care provider values. There is, however, little clarity about how DCEs can be used to assess the concordance of patient and health care provider preferences and provide insights into improving decision quality at the individual level. The aim of this paper is to review studies comparing patient and health care provider preferences for health care interventions using DCEs, specifically to (1) quantify the extent to which they demonstrate concordance of patient and health care provider preferences, and (2) to review the methodology of DCEs to evaluate similarities, differences and rigour of their designs.

2. METHODS

2.1 Systematic search

Search terms were entered into Medline, EMBASE, Econlit, PyscINFO and Web of Science between the dates of July 28th and July 31st, 2015. The search strategy combined free text and MeSH terms pertaining to three main concepts - preferences, patients, and health care practitioners with prescribing authority – with free text and MeSH terms generally describing discrete choice experiments. The final search line was defined as the combinations of groups of terms as follows: (Patients AND Preferences) AND (Prescribers AND Preferences) AND (Discrete choice). An example of the search strategy is presented in the appendix. The search was validated by checking that all references from two previous systematic reviews involving discrete choice experiments[13,20], which reported experimental results from both patient and health care practitioner samples were captured.

After the removal of duplicates, title and abstract screening to select relevant studies was performed; the entire search list was divided and reviewed independently by NB, MH and KM such that every citation was looked at by two reviewers. The inclusion criteria were as follows: peer-reviewed, complete text of an empiric journal article; English language; published between 1995, when relevant DCEs have been published in health[20,21], and present (July 31st, 2015); discussing a health care related topic or condition; eliciting preferences by discrete-choice (DCE) methodology (modified DCE, rankings, adapted conjoint analysis, conjoint analysis were excluded); containing a comparison of patients' or their caregivers' preferences to those of health care practitioners with prescribing authority using the same

DCE questionnaire (i.e. with the same attributes). If a study could not be excluded with certainty, it was included in the full text review. After independent title and abstract review, there was a 96% agreement (see appendix) between all three reviewers, who then met to resolve any disagreement.

2.2. Data Extraction

Data extraction focused on describing the characteristics of the paper based on the key design, analysis and interpretation components of DCEs relevant to this study, identified from a checklist conceptualizing the key components for critical appraisal[22]. These data included methods of designing choices and attributes used in the DCE, piloting, study samples, framing, marginal rates of substitution, and the analysis, including consideration of subgroups and heterogeneity.

We classified attributes in line with the systems-based framework of structures, processes and outcomes approach outlined by Donabedian[23,24] and operationalised in previous reviews in this area[25] for consistency. Structure corresponds to the settings in which health care occurs (including material resources, organisational structures and human resources), process to the factors related to the delivery of care (including the patient's care seeking, and health care practitioners making a diagnosis or recommending a treatment[24]), and outcomes are the effects or consequences of health care or treatment on the patient's overall health status, behaviour and satisfaction with care[24]. We chose to use this framework and classify all attributes from the papers retrieved by the systematic review, in order to ascertain whether the concordance between patient and prescriber preferences varies depending on where a given attribute lies in the health care framework. Any assumptions made during the process of the classification are declared.

Finally, we extracted details of how patient/non-health care practitioner and health care practitioner were assessed for concordance or disagreement, based on the methods reported in each of the studies included in the review and guided by the literature on high-quality decision making[5]. We considered whether the analysis of subgroup or heterogeneity were used to inform this assessment of concordance of preferences.

2.3. Data Synthesis

We attempted to synthesize coefficients derived from each study to observe patterns in attribute types where there was more or less concordance between patients and health care providers. Comparing coefficients from DCEs is challenging and limited by differences in scales where separate DCEs are used in patients and health care providers within each study, and different DCEs between studies.[26] We follow an approach previously used[27,28] where we crudely estimate the relative importance of each attribute (based on the classification described above) by dividing range of coefficients for each attribute by the sum of all coefficient ranges, to provide the rank of importance. We then compared the difference in ranks between the patient and health care provider attributes. Since studies had a different numbers of attributes, we then divided the differences by the number of attributes to provide a score. Finally, we simply took the weighted average of this score by attribute classification.

3. RESULTS

3.1 Systematic review summary

Our search strategy resulted in 1532 hits, from which we retrieved 140 studies after title and abstract review. After full text review, 38 papers were selected to be included in the review (figure 1).

The selected studies came from 15 countries, with five including multiple countries within the same study[29–33]. The majority of studies were from the UK (n=9), the Netherlands (n=7) or Canada (n=5). The studies covered a range of interventions, the most common being drug treatment (n=12), screening (n=8); of which 4 were prenatal screening), provision of services (for example day case surgery[34] or rehabilitation/occupational therapy[35]) and provision of treatment (both for infertility[30,36]). These interventions were based in 26 different indications/diseases, with nine relating to eight different types of cancer[9,29,37–43], three in Down's syndrome[44–46], two relating to kidney disease and organ transplantation[47,48], two in haemophilia[49,50], and two in infertility[30,36] (Table 1).

Table 1: Summary of studies

Country	**N(%) out of 38 studies**
UK	9 (24%)
Netherlands	7 (18%)
USA	3 (8%)
Canada	5 (13%)
Australia	4 (11%)
Italy	4 (11%)
Germany	3 (8%)
Other (for all <n=2)	10 (26%)
Disease	

Cancer	10 (26%)
Chronic Diseases	10 (26%)
Down's Syndrome	3 (8%)
Hemophilia	2 (5%)
Infertility	2 (5%)
Pregnancy and/or Childbirth	2 (5%)
Other (for all <n=2)	9 (24%)
Intervention	
Screening (including prenatal and genetic testing)	9 (24%)
Provision of services (including nondrug treatments – i.e. surgery or occupational therapy)	11 (29%)
Drug Treatment Preferences	16 (42%)
Preferred Symptoms	1 (3%)
Valuation of Health States	1 (3%)
Survey administration	
Self-completed (online)	6 (16%)
Self-completed (postal)	11 (29%)
Self-completed (on location)	6 (16%)
Interview administered	4 (11%)
Mixture of online and postal (one each group)	2 (5%)
Mixture of interview and postal (one each group)	1 (3%)
Mixture of on-location and postal (one each group)	2 (5%)
Mixture of any of the above	7 (18%)
Not reported	None
Number of attributes	**N (%) out of 43 DCEs in 38 studies**
2-4	10 (23%)
5-7	29 (67%)
8-10	4 (9%)
> 10	None
Not reported	None

Attributes covered	**N (%) out of 230**
Structure	19 (8%)
Process	144 (63%)
Outcome	67 (29%)
Development of attributes	**N (%) out of 38 studies**
Focus groups	7 (18%)
Interviews or panel consultations (in one or more groups)	15 (40%)
Mixture of focus group and interviews (one method each group)	6 (16%)
Other methods (including mixed methods)	8 (21%)
Not described	2 (5%)
Development of survey	**N(%) out of 38 studies**
Piloting in all groups of respondents	12 (32%)
Piloting in one (but not at all) groups of respondents	4 (11%)
Piloting in neither group of respondents	3 (5%)
Piloted but group not described/unclear	5 (13%)
Not described	14 (37%)

Notes:

1. Studies taking place in multiple countries were listed under each of the countries separately (total % is more than 100%). Multi-country studies took place in Norway and Denmark[31], United States and Canada[29], The Netherlands and Belgium[30], France, Germany, Italy, Spain and the United Kingdom[33], and Japan and the United States[32].

2: One study required participants to complete the DCE twice, once with assistance and once at home soon after[52]; this was entered as both interview and self-complete via post

3: Five studies[9,33,37,55,65] included separate DCEs for the HCP and non-HCP populations; the numbers of attributes for each DCE were entered independently

3.2 Choice and attribute design and piloting

Of the studies we reviewed that reported the process of survey development, 36 (95%) reported the source of attributes used in their DCE and 24 (63%) reported having piloted their study. The groups that were used to generate attributes and pilot surveys varied. 14 (39%) of the studies that reported their attribute generation sought input from people representative of all groups who would be asked to complete the DCE[29–33,51–59], and 13 (54%) of the studies that reported the piloting in their study piloted the survey in all respondent groups[29,38,40,43,45,48–50,53,54,58–60]. There were only five studies that reported having generated their attributes and piloted their survey in all groups of respondents[29,53,54,58,59]. The method used to generate attributes included focus groups, interviews or panel consultations, a combination of focus groups and interviews, or other methods.

Where studies did not report having generated attributes using input from all respondent groups, there was an equal split between those that generated attributes using only health care providers (n=6)[37,38,40,48,50,61], non-health care providers (n=8)[34,35,42,44,46,54,60,62], or neither (n=7)[9,36,41,43,45,47,63]. Those that reported using neither respondent groups most often used literature reviews, or information from regulatory requirements and product labelling.

Piloting in all target groups was more common (13 of 22 studies reporting piloting), and a number of studies that did not report generating attributes in all groups reported piloting their survey in all respondents[38,40,45,48,50].

3.3. Attribute classification

There were a total of 230 attributes included in the 38 studies, of which 144 (63%) could be classified as process attributes, 67 (29%) as outcomes, and 19 (8%) as structure. The most common attribute type related to delivery and timing (n=57) and morbidity (n=39), followed by safety (n=29), access (n=26), patient/physician relationships (n=20), mortality (n=16), continuity and co-ordination of care (n=13), health-related quality of life (n=11), legal issues (n=4), infrastructure (n=7), financial issues (n=5), and qualifications of health care providers (n=3). Four DCEs, all looking at issues of screening and testing contained only diagnosis and testing attributes[39,44,46,54], and one looking at varying drug effects in diabetes, contained only morbidity attributes[56].

3.4 Study sample and framing

The DCEs identified tried to compare the preferences of multiple different groups, separating out preferences of health care providers and non-health care providers (Table 2). The composition of the studies was heterogeneous; although just over half of the studies (n=20) compared one group of health care providers with either patients (or in the case of screening, those who had experienced the test)[9,30,32,34,36,38–43,45,51,55,57,59,60,62,64] or the general population[29], 12 studies compared the preferences of patients [31,33,35,44,46,47,49,50,53,56,65] or parents of patients[37] with multiple health care providers, and the remainder varied in either their non-health care provider[52,58,61] or both non-health care provider and health care provider groups[48,54,63].

Table 2: Matrix of preferences sought

	Non-health care		
Health care Professionals	Patients	General public	Parents or caregivers
GP	14 (37%)	5 (13%)	4 (11%)
Dentist	1 (3%)	0 (0%)	0 (0%)
Surgeon	2 (5%)	1 (3%)	2 (5%)
Other physician specialty	12 (32%)	4 (11%)	4 (11%)
Nurse/ nurse specialist	6 (16%)	1 (3%)	3 (8%)
Pharmacist	3 (8%)	0 (0%)	3 (8%)
Other Professions	9 (24%)	2 (5 %)	1 (3%)
Health care trainee	1 (3%)	1 (3%)	1 (3%)

Notes: Where papers sought preferences from multiple stakeholders, these are counted individually. Data show % of times each profession was involved in the 38 studies

The framing of the decision given to the health care providers and non-health care providers was almost equally split between being the same instruction for both groups (n=18)[33,34,36,39–41,43,47–53,55,59,61,63], being different for each group (n=19)[9,29–32,35,37,38,42,44–46,54,56–58,60,62,64], or unclear[65].

The dominant framing of the question where the instruction was the same was to pick between the option, with no specific framing of who they were making the decision for reported[33–36,39–41,43,47,48,51–53,55,61,63], although some studies indicated that the health care providers were asked to choose the option with the biggest global benefit, for themselves[50].

Where the instruction was different, the main difference was that the non-health care provider groups made the decision as the patient whilst the health care provider made the decision they would recommend to their patient [9,29–32,37,38,42,44–46,54,57,58,60,62]. One study framed the same vignette about treating a patient in three different ways to different groups, asking regulators which treatment was appropriate, physicians what they would recommend, and patients what they would prefer if they were the patient in the vignette[56]. Two studies took an alternative perspective, asking the health care provider to attempt to predict their patient choice or preference, providing an alternative angle of concordance between patient and physician preferences[42,64]

3.5 Analysis: methods, marginal rates of substitution & sub-groups/heterogeneity

The most common analytical methods used were the random-effects probit[34,41,42,44,46,48,49,53,62,64,65], the logit/conditional logit[31,35,45,54,56,59,63], random effects logit[32,38,50,61], multinomial logit[30,33,37,47,52], or mixed logit[9,40,51,58], with five studies using a range of other methods[29,36,43,55,57].

One or more of the attributes in the DCE was used to scale coefficients of other attributes in 23 studies; most commonly this was monetary (n=9)[29,34–37,39,50,55,57,65], time (n=9) [31,43–46,48,53,54,63], accuracy of testing (n=4)[45,53,54,57], or risk (n=3)[41,44,51]. One study specified they had framed their cost attribute in a different way for patients (out of pocket payment) and physicians (as additional hospital cost)[61].

The majority of studies (n=34) reported accounting for heterogeneity within samples; this was most commonly analysed using sub-groups[30,31,35,36,42,45,48–50,52–54,57,59,61,63,65], or incorporating respondent information as covariates in the model[9,29,41,58,62]. In other studies, heterogeneity was accounted for by allowing random parameters in the model to be estimated [32,38] or using a heteroskedastic condition logit model[59]. One study reported that heterogeneity in preferences existed because the mean co-efficient for a physician was non-significant, but the standard deviation for the point estimate was significant[40]. Only one study explored different subgroups of respondents using latent class analysis, finding two segments of respondents which differed in their order of preferences but could not be differentiated by their characteristics[55].

3.6. Approaches to measuring concordance

There was no consistent approach to measuring the concordance of preferences between health care providers and non-health care providers. The methods used varied widely, and could be grouped into three broad approaches for descriptive purposes; (1) qualitative comparison of regression coefficients, (2) statistical tests of differences or similarities of coefficients, and (3) regression diagnostics. Although there was variation in the methods used to assess concordance, all approaches were based on comparisons at the aggregate level of the sample. One study segmented respondents using latent class analysis[55] but did not explore concordance of patients and physicians using this approach.

Where coefficients were compared, a method of ranking of attributes based on the strength of coefficients[9,33,35–38,40–47,54,59,62,64,65] was the most frequently used, although there were approaches based on the difference between coefficients themselves, or the confidence intervals of coefficients[30,49]. An alternative approach to comparing results was to first estimate coefficients on a common scale, using marginal rates of substitution (e.g. using a payment vehicle like willingness to pay[34,36,55], willingness to accept risk[32,51], or time[31]), relative importance[50,61], or some weighting of the model coefficients[58], and then making a comparison of these. One study calculated a ratio of the importance of patient and physician coefficients[32].

There were examples where statistical tests were used to compare the coefficients generated by each of the different groups. These either took the form of correlation-based approaches to look for statistical similarities of coefficients (e.g. Kendall's Tau b[52] or Spearman's rank correlation[41]), or unpaired tests for differences in coefficients between groups (e.g. Pearson-chi square[39]) or pooled regression approaches which look for statistical differences between groups on coefficients.[60,62]

A final set of tests to assess differences in preferences was regression based diagnostics. These included the use of interaction terms or the Wald test, to see if pooled analysis with a respondent group identifier or interaction term was statistically significant[60] or significantly improved model fit[29,32,48,65]. An alternative approach tested whether the coefficients in two regressions analyses using different data sets were equal (Chow test) [56]. A third approach was to assess the impact on the scale parameters of different datasets (Swait & Louviere test) [51,53,59] to see whether models needed to be estimated separately if there were underlying differences in the two datasets.

3.7 Findings of concordance

The conclusions from the studies (Table 3) discussed mixed concordance and discordance most frequently (28 studies), followed by discordance of patient and health care provider preferences (12 studies) and concordant preferences (4 studies). These conclusions appeared to be fairly consistent irrespective of the methods used to test for concordance.

Table 3: Summary of concordance analysis and resulting conclusions

		Author conclusion		
Method used	N(%)	Evidence of concordance N (%)	Evidence of disagreement N (%)	Mixed (N%)
Qualitative comparison				
Strength of coefficients	19 (50%)	2 (11%)	3 (16%)	14 (74%)
MRS	6 (16%)	-	3 (50%)	3 (50%)
Relative importance	2 (5%)	-	-	2 (100%)
Weighting	1 (3%)	-	-	1 (100%)

Difference	2 (5%)	-	1 (50%)	1 (50%)
Statistical tests				
Similarity	2 (5%)	-	-	2 (100%)
Unpaired differences	1 (3%)	-	-	1 (100%)
Pooled regression	2 (5%)	-	2 (100%)	-
Regression diagnostics				
Wald test/interactions	5 (13%)	-	2 (40%)	3 (60%)
Chow	1 (3%)	1 (100%)	-	-
Swait & Louviere test	3 (8%)	1 (33%)	1 (33%)	1 (33%)

Synthesis of the reported coefficients from the studies reviewed showed that concordance and discordance varied by the type and classification of attribute (figure 2). Health care professionals appeared to believe that structure and outcome attributes were of greater importance than patients did, whereas patients appeared to place greater importance on process outcomes than health care professionals. However, within these categories results differed by attribute.

The greatest discordance between patients and health care providers were for mortality (e.g. “Risk of death”[37]) whereby health care providers believed it to be more important than patients, and infrastructure (e.g. “Number of beds in hospital room”[62]) whereby health care providers again thought that this was more important than patients .

Patients placed more importance on issues of process, for example safety (e.g. “Risk of urinary incontinence due to treatment”[40]) than health care providers. Similarly, patients cared more about delivery and timing (e.g. “Route of drug administration”[60]) than health care providers. There was evidence of some discordance around issues of patient and health care provider relationship (e.g. “Physician’s attitude to patients”[30]), morbidity (e.g. “Time necessary to recover (defined as returning to normal activities)”[61]) and access (e.g. “Increase in health care taxes (cost)”[50]) with patients rating this as more important than health care providers.

Finally, there was good concordance between patients and health care providers on aspects of control and coordination (e.g. “Continuity of physicians”[30]) and quality of life (e.g. “Quality of life”[41]).

4. DISCUSSION

This systematic review found that discordance between patient and health care provider preferences in decisions around health care interventions appears to be more common than concordance, and secondly that concordance (or discordance) is not a binary concept. We identified a large body of work which has attempted to understand the similarity or differences of patient and health care provider preferences. Most of these studies have reported mixed conclusions on the concordance of preferences for patients and health care providers, but there is more evidence of discordance than concordance. Similarly, we found that concordance or discordance of patient and health care professional preferences, in the DCE context, varies across the different attributes being considered. In addition, concordance seems to differ according to the individuals involved in making the choices. However, the studies did not provide results to allow us to understand whether and how the importance of different attributes varies within these groups of respondents, which is a limitation of the literature we reviewed.

One of the key limitations of the literature we reviewed was that the reasons for differences in preferences between patients and health care providers were unclear. We found considerable variation in the approaches used both between and within the DCEs we reviewed, including methods of analysis and testing of differences or variation in preferences. Almost half of the studies we reviewed used different versions of the survey in patients and health care provider groups, meaning that differences in preferences between groups could potentially be attributable to genuine differences, or alternatively to differences in the choice sets they completed. As we have noted, where DCEs which differ in the framing are used, or the attributes or levels included, comparison of coefficients for the purpose of assessing concordance of preferences is challenging and limited by differences in scales. For this reason, we recommend that any DCE aimed at assessing concordance should use the same attributes and levels, and should report the framing of the question to allow readers to assess whether the tasks are equivalent.

Comparisons of patients and health care provider preferences in the DCEs were also primarily made at the aggregate level, which is not informative about level of agreement and distribution of sub-group preferences. One study used a latent class analysis approach to identify sub-group preferences within patient and health care provider preferences, but did not use this approach to understand whether combinations of patient and health care provider sub-groups had more concordant preferences[55]. There is a need to try to identify groups of patients and health care providers with similar preferences using latent class methods. Within groups of patients and health care providers there are likely to be subgroups which are fairly homogenous in their preferences for certain aspects of treatment, but these preferences may be heterogeneous to those of other groups of patients and health care providers. We recommend that studies seeking to understand the concordance of patient and provider preferences should investigate the heterogeneity of preferences within groups and how these relate. Knowledge of the existence of these groups and their preferences will help inform whether the same treatments, programs or services can be offered to everyone, whether different options can be offered to different groups, or whether patients and health care providers can be more closely matched according to their preferences. Failing to account for heterogeneity in patient and health care provider preferences might mean that a treatment or service could be designed which meets the preferences of the aggregate group, but fails to meet the preferences of sub-groups of people within that population[66].

This study is not without limitations. Firstly, by restricting ourselves to studies using DCE methodology, we took a narrow view of the literature on this topic. However, assessing the consistency of preferences across multiple different methodologies would have introduced considerable additional heterogeneity, making it even more difficult to draw firm conclusions. Secondly, the way we synthesized coefficients between and across studies required assumptions that are known to be problematic. However, we believe that the value of undertaking a synthesis of results adds to the narrative review of the literature in highlighting the key areas where concordance and discordance is greatest based on the current evidence. For this reason, we believe it is worth making these assumptions, but recognize that this limits the interpretation of this synthesis. Additionally, the terms we used in our literature search strategy meant that, in some cases, papers from studies which set out to compare patient or health care provider preferences, but reported these in separate publications based on a single sample of respondents would not have been incorporated into the review. Finally, it is unclear what the concordance (or lack thereof) of preferences or values means within each type of attribute as regards the patient-centeredness of care or the quality of the treatment decision.

DCE studies provide an excellent opportunity for determining whether there is concordance of values and preferences for aspects of treatments or services between patients and providers. However, our findings highlight that no consistent approach has been taken to understand whether there is concordance, and we have identified a number of issues which have limited the interpretation of the approaches we identified and made recommendations for future studies.

We have also shown that discordance in patient and health care provider preferences appears to be common, and that concordance (or discordance) varies according to which attributes are being considered. For example, for a single decision there could be concordance on the importance of quality of life outcomes, but discordance around mortality outcomes and issues of access. This highlights that concordance should not been considered as a binary outcome and it is important that any measure of value concordance considers all aspects jointly.

References

1. Stewart, M. Towards a global definition of patient centred care. *BMJ* **322,** 444–445 (2001).
2. Little, P. *et al.* Preferences of patients for patient centred approach to consultation in primary care: observational study. *BMJ* **322,** 468 (2001).
3. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* **268,** 2420–2425 (1992).
4. Haynes, R. B., Devereaux, P. J. & Guyatt, G. H. Physicians' and patients' choices in evidence based practice. *BMJ* **324,** 1350 (2002).
5. Elwyn, G. *et al.* Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* **333,** 417 (2006).
6. Mulley, A., Trimble, C. & Elwyn, G. The problem of silent misdiagnosis. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/patients-preferences-matter-may-2012.pdf. (Accessed: 23rd August 2016)
7. Mulley, A. G., Trimble, C. & Elwyn, G. Stop the silent misdiagnosis: patients' preferences matter. *BMJ* **345,** e6572 (2012).
8. Stacey, D. *et al.* Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst. Rev.* CD001431 (2011). doi:10.1002/14651858.CD001431.pub3
9. Fiebig, D. G., Haas, M., Hossain, I., Street, D. J. & Viney, R. Decisions about Pap tests: What influences women and providers? *Soc. Sci. Med.* **68,** 1766–1774 (2009).
10. Jalukar, V., Funk, G. F., Christensen, A. J., Karnell, L. H. & Moran, P. J. Health states following head and neck cancer treatment: patient, health-care professional, and public perspectives. *Head Neck* **20,** 600–608 (1998).

11. Sassi, F., Mcdaid, D. & Ricciardi, W. Conjoint analysis of preferences for cardiac risk assessment in primary care. *Int. J. Technol. Assess. Health Care* **21,** 211–8 (2005).
12. Casparie, A. F. & van der Waal, M. A. Differences in preferences between diabetic patients and diabetologists regarding quality of care: a matter of continuity and efficiency of care? *Diabet. Med. J. Br. Diabet. Assoc.* **12,** 828–832 (1995).
13. Axel C. Mühlbacher & Christin Juhnke. Patient Preferences Versus Physicians' Judgement: Does it Make a Difference in Healthcare Decision Making? *Patient Prefer. Physicians' Judgement Does It Make Differ. Healthc. Decis. Mak.* **11,** 163–180 (2013).
14. van Helvoort-Postulart, D. *et al.* Investigating the complementary value of discrete choice experiments for the evaluation of barriers and facilitators in implementation research: a questionnaire survey. *Implement. Sci.* **4,** 10 (2009).
15. de Bekker-Grob, E. W., Ryan, M. & Gerard, K. Discrete choice experiments in health economics: a review of the literature. *Health Econ.* **21,** 145–172 (2012).
16. Gold, M. R. *Cost-effectiveness in health and medicine*. (Oxford University Press, 1996).
17. Elliott, R. & Payne, K. *Essentials of Economic Evaluation in Healthcare*. (Pharmaceutical Press, 2005).
18. Reed Johnson, F. *et al.* Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health* **16,** 3–13 (2013).
19. Bridges, J. F. P., Onukwugha, E., Johnson, F. R. & Hauber, A. B. Patient Preference Methods - A Patient Centered Evaluation Paradigm. *ISPOR Connect.* **December, 15,** 4–7 (2007).
20. Harrison, M. *et al.* Risk as an Attribute in Discrete Choice Experiments: A Systematic Review of the Literature. *Patient* **7,** (2014).
21. Ryan, M. & Gerard, K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl. Health Econ. Health Policy.* **2,** 55–64 (2003).

22. Lancsar, E. & Louviere, J. Conducting discrete choice experiments to inform Healthcare decision making. *Pharmacoeconomics* **26,** 661–677 (2008).

23. Donabedian, A. Evaluating the Quality of Medical Care. *Milbank Mem. Fund Q.* **44,** 166–206 (1966).

24. Donabedian A. The quality of care: How can it be assessed? *JAMA* **260,** 1743–1748 (1988).

25. Mühlbacher, A. C. & Juhnke, C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? *Appl. Health Econ. Health Policy.* **11,** 163–180 (2013).

26. Flynn, T. N., Louviere, J. J., Peters, T. J. & Coast, J. Using discrete choice experiments to understand preferences for quality of life. Variance-scale heterogeneity matters. *Soc. Sci. Med.* **70,** 1957–1965 (2010).

27. Vick, S. & Scott, A. Agency in health care. Examining patients' preferences for attributes of the doctor–patient relationship. *J. Health Econ.* **17,** 587–605 (1998).

28. Malhotra, C., Chan, N., Zhou, J., Dalager, H. B. & Finkelstein, E. Variation in physician recommendations, knowledge and perceived roles regarding provision of end-of-life care. *BMC Palliat. Care* **14,** (2015).

29. Marshall, D. A. *et al.* How do physician assessments of patient preferences for colorectal cancer screening tests differ from actual preferences? A comparison in Canada and the United States using a stated-choice survey. *Health Econ.* **18,** 1420–1439 (2009).

30. Empel, I. W. H. van *et al.* Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. *Hum. Reprod.* **26,** 584–593 (2011).

31. Espelid, I. *et al.* Preferences over dental restorative materials among young patients and dental professionals. *Eur. J. Oral Sci.* **114,** 15–21 (2006).

32. Okumura, K. *et al.* Comparing Patient and Physician Risk Tolerance for Bleeding Events Associated with Anticoagulants in Atrial Fibrillation—evidence from the United States and Japan. *Value Health Reg. Issues* **6,** 65–72 (2015).

33. Chancellor, J., Martin, M., Liedgens, H., Baker, M. G. & Müller-Schwefe, G. H. H. Stated Preferences of Physicians and Chronic Pain Sufferers in the Use of Classic Strong Opioids. *Value Health* **15,** 106–117 (2012).

34. Gidman, W., Elliott, R., Payne, K., Meakin, G. H. & Moore, J. A comparison of parents and pediatric anesthesiologists' preferences for attributes of child daycase surgery: a discrete choice experiment. *Pediatr. Anesth.* **17,** 1043–1052 (2007).

35. Laver, K., Ratcliffe, J., George, S., Lester, L. & Crotty, M. Preferences for rehabilitation service delivery: A comparison of the views of patients, occupational therapists and other rehabilitation clinicians using a discrete choice experiment. *Aust. Occup. Ther. J.* **60,** 93–100 (2013).

36. Huppelschoten, A. G. *et al.* The monetary value of patient-centred care: results from a discrete choice experiment in Dutch fertility care. *Hum. Reprod.* **29,** 1712–1720 (2014).

37. Regier, D. A. *et al.* Discrete Choice Experiment to Evaluate Factors That Influence Preferences for Antibiotic Prophylaxis in Pediatric Oncology. *PLoS ONE* **7,** (2012).

38. Gatta, F. *et al.* Patients' and Physicians' Preferences for Approaches to Bone Metastases Treatment in Turkey. *Int. J. Hematol. Oncol.* **25,** 118–129 (2015).

39. Boone, D. *et al.* Patients' & Healthcare Professionals' Values Regarding True- & False-Positive Diagnosis when Colorectal Cancer Screening by CT Colonography: Discrete Choice Experiment. *PLoS ONE* **8,** e80767 (2013).

40. de Bekker-Grob, E. W. *et al.* Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. *Br. J. Cancer* **109,** 633–640 (2013).

41. Thrumurthy, S. G., Morris, J. J. A., Mughal, M. M. & Ward, J. B. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric cancer. *Br. J. Surg.* **98,** 1124–1131 (2011).

42. Mühlbacher, A. C. & Nübling, M. Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. *Eur. J. Health Econ.* **12,** 193–203 (2011).

43. Shafey, M., Lupichuk, S. M., Do, T., Owen, C. & Stewart, D. A. Preferences of patients and physicians concerning treatment options for relapsed follicular lymphoma: a discrete choice experiment. *Bone Marrow Transplant.* **46,** 962–969 (2011).

44. Bishop, A. J. *et al.* Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. *BJOG Int. J. Obstet. Gynaecol.* **111,** 775–779 (2004).

45. Hill, M., Fisher, J., Chitty, L. S. & Morris, S. Women's and health professionals' preferences for prenatal tests for Down syndrome: a discrete choice experiment to contrast noninvasive prenatal diagnosis with current invasive tests. *Genet. Med.* **14,** 905–913 (2012).

46. Lewis, S. M., Cullinane, F. M., Carlin, J. B. & Halliday, J. L. Women's and health professionals' preferences for prenatal testing for Down syndrome in Australia. *Aust. N. Z. J. Obstet. Gynaecol.* **46,** 205–211 (2006).

47. Davison, S. N., Kromm, S. K. & Currie, G. R. Patient and health professional preferences for organ allocation and procurement, end-of-life care and organization of care for patients with chronic kidney disease using a discrete choice experiment. *Nephrol. Dial. Transplant.* **25,** 2334–2341 (2010).

48. Clark, M. D. *et al.* Who should be prioritized for renal transplantation?: Analysis of key stakeholder preferences using discrete choice experiments. *BMC Nephrol.* **13,** 152 (2012).

49. Mantovani, L. G. *et al.* Differences between patients', physicians' and pharmacists' preferences for treatment products in haemophilia: a discrete choice experiment. *Haemophilia* **11,** 589–597 (2005).

50. Scalone, L., Mantovani, L. G., Borghetti, F., Von Mackensen, S. & Gringeri, A. Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA Study. *Haemophilia* **15,** 473–486 (2009).

51. Johnson, F. R. *et al.* Are Gastroenterologists Less Tolerant of Treatment Risks than Patients? Benefit-Risk Preferences in Crohn's Disease Management. *J. Manag. Care Pharm.* **16,** 616–628 (2010).
52. Bijlenga, D., Birnie, E., Mol, B. W. & Bonsel, G. J. Obstetrical outcome valuations by patients, professionals, and laypersons: differences within and between groups using three valuation methods. *BMC Pregnancy Childbirth* **11,** 93 (2011).
53. Payne, K. *et al.* Valuing pharmacogenetic testing services: A comparison of patients' and health care professionals' preferences. *Value Health* **14,** 121–134 (2011).
54. Hill, M., Suri, R., Nash, E. F., Morris, S. & Chitty, L. S. Preferences for Prenatal Tests for Cystic Fibrosis: A Discrete Choice Experiment to Compare the Views of Adult Patients, Carriers of Cystic Fibrosis and Health Professionals. *J. Clin. Med.* **3,** 176 (2014).
55. Deal, K., Keshavjee, K., Troyan, S., Kyba, R. & Holbrook, A. M. Physician and patient willingness to pay for electronic cardiovascular disease management. *Int. J. Med. Inf.* **83,** 517–528 (2014).
56. Mol, P. G. M. *et al.* Understanding drug preferences, different perspectives. *Br. J. Clin. Pharmacol.* **79,** 978–987 (2015).
57. Beulen, L. *et al.* Women's and healthcare professionals' preferences for prenatal testing: a discrete choice experiment. *Prenat. Diagn.* **35,** 549–557 (2015).
58. Whitty, J. A. *et al.* Preferences for the delivery of community pharmacy services to help manage chronic conditions. *Res. Soc. Adm. Pharm.* **11,** 197–215 (2015).
59. Pedersen, L. B., Kjær, T., Kragstrup, J. & Gyrd-Hansen, D. Do General Practitioners Know Patients' Preferences? An Empirical Study on the Agency Relationship at an Aggregate Level Using a Discrete Choice Experiment. *Value Health* **15,** 514–523 (2012).
60. de Bekker-Grob, E. W., Essink-Bot, M.-L., Meerding, W. J., Koes, B. W. & Steyerberg, E. W. Preferences of GPs and Patients for Preventive Osteoporosis Drug Treatment - Springer. *Pharmacoeconomics* **27,** 211–219 (2009).

61. Faggioli, G., Scalone, L., Mantovani, L. G., Borghetti, F. & Stella, A. Preferences of Patients, Their Family Caregivers and Vascular Surgeons in the Choice of Abdominal Aortic Aneurysms Treatment Options: The PREFER Study. *Eur. J. Vasc. Endovasc. Surg.* **42,** 26–34 (2011).

62. Neuman, E. & Neuman, S. Agency in health-care: are medical care-givers perfect agents? *Appl. Econ. Lett.* **16,** 1355–1360 (2009).

63. Park, M.-H., Jo, C., Bae, E. Y. & 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* **15,** 933–939 (2012).

64. Mühlbacher, A. C., Stoll, M., Mahlich, J. & Nübling, M. Evaluating the concordance of physician judgments and patient preferences on AIDS/HIV therapy - a Discrete Choice Experiment. *Health Econ. Rev.* **3,** 1–11 (2013).

65. Lee, A., Gin, T., Lau, A. S. C. & Ng, F. F. A comparison of patients' and health care professionals' preferences for symptoms during immediate postoperative recovery and the management of postoperative nausea and vomiting. *Anesth. Analg.* **100,** 87–93 (2005).

66. Deal, K. Segmenting Patients and Physicians Using Preferences from Discrete Choice Experiments. *Patient* **7,** 5–21 (2014).

Figure 1: PRISMA Flow Diagram for Systematic Review

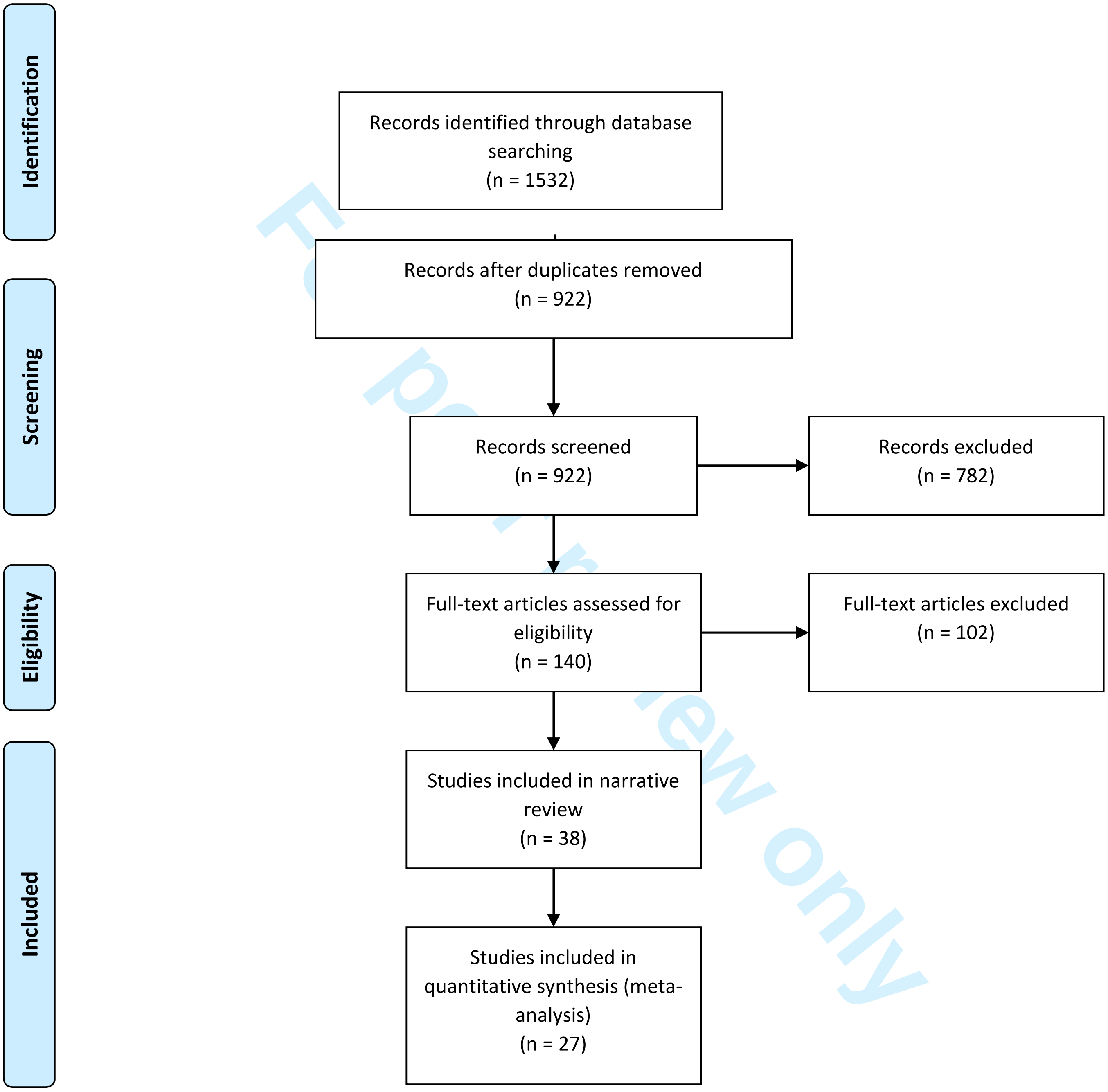
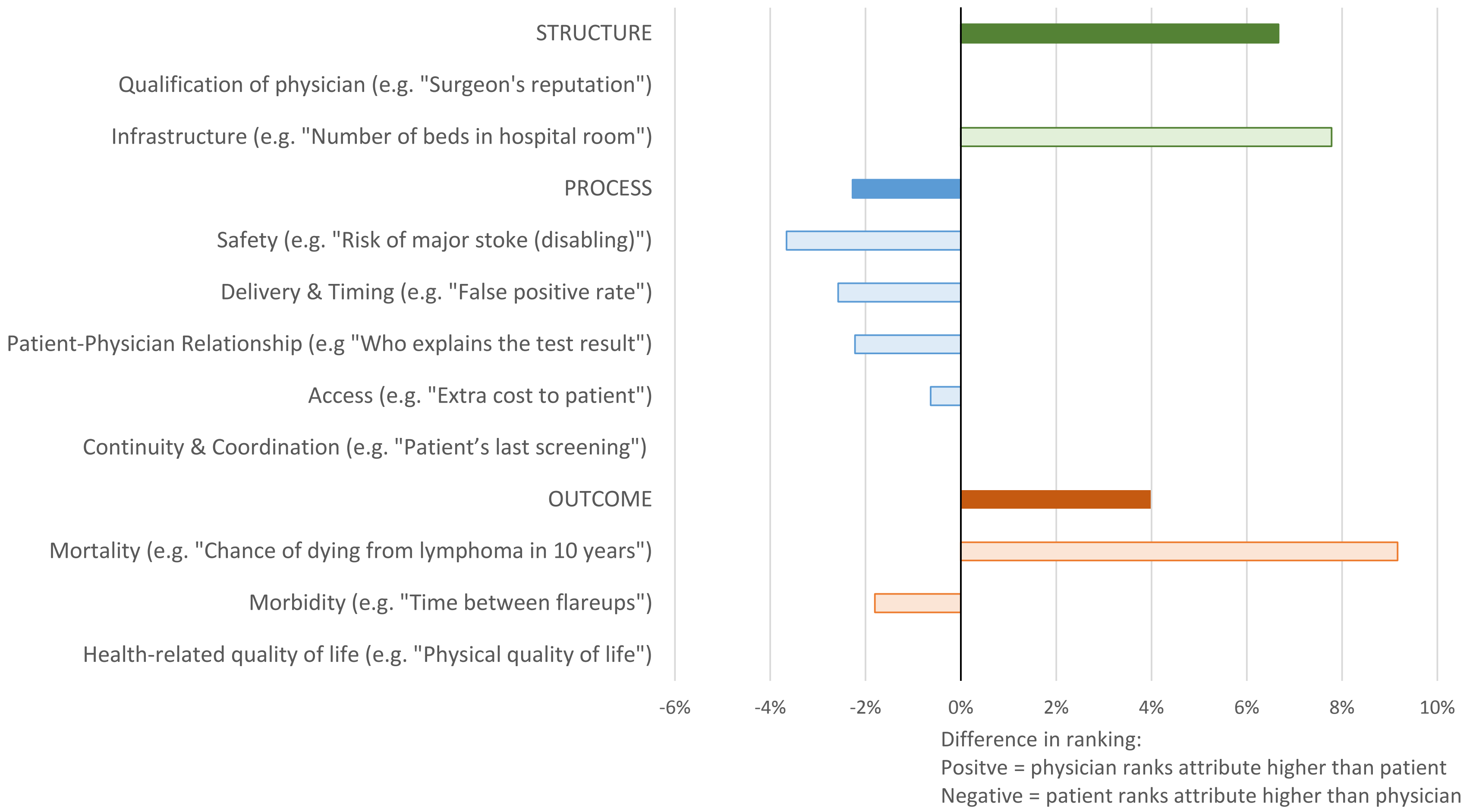


Figure 2: Synthesis of concordance between patients and physician preferences for different types of attributes



Chance of dying from lymphoma in 10 years[51] ; "Number of beds in hospital room"[62] ; Time between flareups[51]; Risk of major stoke (disabling)[32]; False positive rate[57]; Who explains the test result[53]; Extra cost to patient[65]; Patient's last screening[9]; Physical quality of life[42]; Reputation of surgeon[41]

Appendix 1: Search Terms: MEDLINE July 28th

Patients

patient*.mp OR Patients/ OR women.mp OR men.mp OR parent*.mp OR caregiver.mp OR carer*.mp

Prescribers

prescriber*.mp OR physician*.mp OR provider*.mp OR expert*.mp OR doctor*.mp OR specialist*.mp OR professional*.mp OR surgeon*.mp

Preference

preference*.mp OR value*.mp OR priorit*.mp OR perspective*.mp

Discrete-choice experiment

part-worth utility*.mp OR paired comparison*.mp OR pairwise choice*.mp OR stated preference.mp OR (discrete adj choice$).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] OR (conjoint adj (analys$ or measurement$ or study or studies or choice$)).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

Final Search:

(Patients AND **Preference)** AND **(Prescribers** AND **Preference)** AND **(Discrete-choice experiment)**

Appendix 2: Agreement statistics

Screening decision		Agreement Overall n	Reviewer 2 n	Reviewer 3 n
Reviewer 1	Reviewer 2 or 3	922	450	473
Y	Y	68	28	40
Y	N	19	10	9
N	N	820	399	421
N	Y	15	12	3
	N agree	888	427	461
	Agreement %	96%	95%	97%

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6, Appendix
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7, Appendix
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7,8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7,8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-11, Tbl. 1-2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-13
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Figure 2
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8-14,16
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	11-13, Figure 2
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	14
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	14-15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	4

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of Discrete Choice Experiments

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014719.R1
Article Type:	Research
Date Submitted by the Author:	09-Feb-2017
Complete List of Authors:	Harrison, Mark; University of British Columbia, Pharmaceutical Sciences; UBC Milbers, Katherine; Centre for Health Evaluation and Outcome Sciences Hudson, Marie; Jewish General Hospital, Lady Davis Institute for Medical Research Bansback, Nick; School of Population and Public Health, University of British Columbia
Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Health services research, Research methods, Health economics
Keywords:	Discrete Choice Experiment, Health Care Preferences, Shared Decision-Making

Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of Discrete Choice Experiments

Mark Harrison[1,2], Katherine Milbers[2], Marie Hudson[3,4,5], Nick Bansback[2,6]

[1] Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC

[2] Centre for Health Evaluation and Outcome Sciences, St Paul's Hospital, Vancouver, BC

[3] Department of Medicine, McGill University, Montréal, QC

[4] Division of Rheumatology, Jewish General Hospital, Montréal, QC

[5] Lady Davis Institute for Medical Research, Montréal, QC

[6] School of Population and Public Health, University of British Columbia, Vancouver, BC

Correspondence to:

Mark Harrison,

Faculty of Pharmaceutical Sciences,

University of British Columbia,

2405 Wesbrook Mall

Vancouver,

BC Canada V6T 1Z3

+1 604-827-0687

mark.harrison@ubc.ca

Objectives: To review studies eliciting patient and health care provider preferences for health care interventions using Discrete Choice Experiments (DCEs) to (1) review the methodology to evaluate similarities, differences, rigor of designs, and whether comparisons are made at the aggregate level or account for individual heterogeneity; and (2) quantify the extent to which they demonstrate concordance of patient and health care provider preferences.

Methods: A systematic review searching Medline, EMBASE, Econlit, PsycINFO and Web of Science for DCEs using patient and healthcare providers. Inclusion criteria: peer-reviewed; complete empiric text in English from 1995-July 31st 2015; discussing a healthcare-related topic; DCE methodology; comparing patients and health care providers.

Design: Systematic review

Results: We identified 38 papers exploring 16 interventions in 26 diseases/indications. Methods to analyze results, determine concordance between patient and physician values, and explore heterogeneity varied considerably between studies. The majority of studies we reviewed found more evidence of mixed concordance and discordance (n=28) or discordance of patient and health care provider preferences (n=12) than of concordant preferences (n=4). A synthesis of concordance suggested that health care providers rank structure and outcome attributes more highly than patients, whilst patients rank process attributes more highly than health care providers.

Conclusions: Discordant patient and health care provider preferences for different attributes of health care interventions are common. Concordance varies according to whether attributes are processes, structures or outcomes, and therefore determining preference concordance should consider all aspects jointly and not a binary outcome. DCE studies provide excellent opportunities to assess value concordance between patients and providers, but assessment of concordance was limited by a lack of consistency in the approaches used and consideration of heterogeneity of preferences. Future DCEs assessing concordance should fully report the framing of the questions, and investigate the heterogeneity of preferences within groups and how these compare.

Strengths and limitations of this study

- By looking specifically at studies using DCE methodology we are using a method which enables preferences and trade-offs in decision making to be understood, but we acknowledge we are taking a focussed view of the literature on this topic.
- We have systematically reviewed a large body of work, which has attempted to understand the similarities or differences of patient and health care provider preferences using DCEs.
- We highlight a lack of consistency within and between studies, which adds difficulty to summarizing findings, but leads to recommendations for future studies.
- We synthesise concordance scores between and across studies to give a sense of differences in concordance by attribute, however the synthesis of concordance scores between and across studies requires assumptions that are known to be problematic.

Funding

This work was supported by a grant from the Canadian Rheumatology Association through the Canadian Initiative for Outcomes in Rheumatology Care (CIORA).

ICME statement

MHa and NB were responsible for substantial contributions to the conception or design of the work, the analysis and interpretation of data, and drafting and critically revising the work. MHu was responsible for substantial contributions to the conception or design of the work and critically revising the work for important intellectual content. KM was responsible for substantial contributions to the acquisition of data for the work and drafting of the manuscript. All authors gave their final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We are grateful to Julia Kaal, University of British Columbia, for her assistance in revising the final draft for publication.

Data sharing statement

This systematic review draws on published research only. No new data were collected or generated in undertaking this study.

Competing interests

Mark Harrison holds the UBC Professorship in Sustainable Health Care, which is funded by Amgen Canada, AstraZeneca Canada, Eli Lilly Canada, GlaxoSmithKline, Merck Canada, Novartis Pharmaceuticals Canada, Pfizer Canada, Boehringer Ingelheim (Canada), Hoffman-La Roche, LifeScan Canada, and Lundbeck Canada.

1. BACKGROUND

Patient-centered care has been long and widely advocated[1,2]. Respectful and responsive to the clinical state, circumstances, preferences and actions of the patient[3,4], it posits clinical decisions to be guided by the values of the patient. Shared-decision making has been promoted as a way to achieve a patient-centred approach by ensuring that patients are fully informed about, and consequently receive options that reflect, their personal treatment preferences[5]. The assumptions of health care providers about their patients' preferences has been termed 'preference diagnosis', and is thought to be as important as the 'medical diagnosis'[6]. Both types of diagnoses involve inferences based on imperfect information; however, while the medical diagnosis is based on a combination of tests, imaging and specialist opinions, it is less clear what informs preference diagnosis. Furthermore, the importance and complexity of diagnosing patient preference is often overlooked, and evidence suggests that health care providers erroneously deem themselves good at diagnosing their patients' preferences[7]. Discordance between patient and health care provider preferences with regard to treatment decisions can lead to preference misdiagnosis, also called the silent misdiagnosis[6].

Preference misdiagnosis affects patients and health systems, as evidence suggest that patients whose care more closely matches their preferences consume less health care[8]. However, evidence is not clear cut on how and to what extent patients and health care providers have different preferences. Studies attempting to understand how and how often patient and health care provider preferences differ have shown mixed results. For example, some studies suggest no differences between patient and health care provider in terms of rank or strength of preference[9] for features of treatments, while others show similar ranks but differences in terms of strength of preference[10], or differences in terms of both rank and strength of preferences[11,12]. A recent review assessing differences between patient and health care provider preferences in health care decision-making found that, at the aggregate level, patients and health care providers had different preferences[13]. However, the evidence was limited and assessment of the preference discordance remains unclear and complicated due to different methodologies (e.g., time trade-off[10], conjoint analysis[11], paired comparison), different disease contexts (e.g., pap tests[9], diabetes[12], head/neck cancer[10], cardiac risk assessment[11]), and types of decision [13]. Furthermore, the conclusions drawn by these studies were comparing aggregate patient and health care provider preferences, which could disguise preference heterogeneity and miss important subgroups with different preferences.

Discrete choice experiments (DCEs) have become an established tool in economic evaluation and decision making[14] and for understanding preferences and predicting choices[14,15] due to their ability to break down and value different components of treatments and services (whether these are processes, structures or outcomes[16–18]) as well as identify the trade-offs people make between these different components[19,20]. While DCEs ask about hypothetical choices, their agreement with actual choices has been shown to be good in a limited number of studies[21,22], although other studies have found conflicting patterns of choices[23], and further evidence of the agreement of hypothetical and revealed (actual) choices is sought[16]. Nevertheless, DCEs theoretically facilitate a realistic assessment of the concordance of patient and health care provider values. There is, however, little clarity about how DCEs can be used to assess the concordance of patient and health care provider preferences and provide insights into improving decision quality at the individual level. The aim of this paper is to review studies which elicit both patient and health care provider preferences for health care interventions using DCEs, specifically to (1) review the methodology of DCEs to evaluate similarities, differences and rigor of their designs, specifically whether comparisons are made at the aggregate level or account for individual heterogeneity; and (2) quantify the extent to which they demonstrate concordance of patient and health care provider preferences.

2. METHODS

2.1 Systematic search

Search terms were entered into Medline, EMBASE, Econlit, PyscINFO and Web of Science between the dates of July 28th and July 31st, 2015. The search strategy combined free text and MeSH terms pertaining to three main concepts - preferences, patients, and health care practitioners with prescribing authority – with free text and MeSH terms generally describing discrete choice experiments. The final search line was defined as the combinations of groups of terms as follows: (Patients AND Preferences) AND (Prescribers AND Preferences) AND (Discrete choice). An example of the search strategy is presented in appendix 1. The search was validated by checking that all references from two previous systematic reviews involving discrete choice experiments[13,16], which reported experimental results from both patient and health care practitioner samples were captured.

After the removal of duplicates, title and abstract screening to select relevant studies was performed; the entire search list was divided and reviewed independently by NB, MH and KM such that every citation was looked at by two reviewers. The inclusion criteria were as follows: peer-reviewed, complete text of an empiric journal article; English language; published between 1995, when relevant DCEs have been published in health[24,25], and present (July 31st, 2015); discussing a health care related topic or condition; eliciting preferences by discrete-choice (DCE) methodology (modified DCE, rankings, adapted conjoint analysis, conjoint analysis were excluded); containing a comparison of patients' or their caregivers' preferences to those of health care practitioners with prescribing authority using the same DCE questionnaire (i.e. with the same attributes). If a study could not be excluded with certainty, it was included in the full text review. After independent title and abstract review, there was a 96% agreement (see appendix 2) between all three reviewers, who then met to resolve any disagreement.

2.2. Data Extraction

Data extraction focused on describing the characteristics of the paper based on the key design, analysis and interpretation components of DCEs relevant to this study, identified from a checklist conceptualizing the key components for critical appraisal[14]. These data included methods of designing choices and attributes used in the DCE, piloting, study samples, framing, marginal rates of substitution, and the analysis, including consideration of subgroups and heterogeneity.

We classified attributes in line with the systems-based framework of structures, processes and outcomes approach outlined by Donabedian[26] and operationalised in previous reviews in this area[27] for consistency. Structure corresponds to the settings in which health care occurs (including material resources, organisational structures and human resources), process to the factors related to the delivery of care (including the patient's care seeking, and health care practitioners making a diagnosis or recommending a treatment[26]), and outcomes are the effects or consequences of health care or treatment on the patient's overall health status, behaviour and satisfaction with care[26]. We chose to use this framework and classify all attributes from the papers retrieved by the systematic review, in order to ascertain whether the concordance between patient and prescriber preferences varies depending on where a given attribute lies in the health care framework. Any assumptions made during the process of the classification are declared.

Finally, we extracted details of how patient/non-health care practitioner and health care practitioner preferences were assessed for concordance or disagreement, based on the methods reported in each of the studies included in the review and guided by the literature on high-quality decision making[5]. We considered whether the analysis of subgroup or heterogeneity were used to inform this assessment of concordance of preferences.

2.3. Data Synthesis

We attempted to synthesize coefficients derived from each study to observe patterns in attribute types where there was more or less concordance between patients and health care providers by developing a concordance score. Comparing coefficients from DCEs is challenging and limited by differences in the variance scale where separate DCEs are used in patients and health care providers within each study, and different DCEs between studies.[28] We follow an approach previously used[29,30] where we crudely estimate the relative importance of each attribute (based on the classification described above) by dividing the range of coefficients for each attribute by the sum of all coefficient ranges within a DCE, to provide the rank of importance of the attribute within that study. We then compared the difference in the rank of importance for an attribute between patient and health care providers. Since different studies have different numbers of attributes, we then divided the differences in the rank of importance of an attribute by the number of other attributes within the DCE to provide concordance score on a common scale (where 0 = perfect concordance of rank importance, -1 indicates that the patient ranks the attribute that the health care provider believes is most important, as the least important, +1 indicates that a health care provider ranks the attribute that the patient believes is most important, as the least important). Finally, we simply took the weighted average of this score across all studies by attribute classification and present these in a figure.

3. RESULTS

3.1 Systematic review summary

Our search strategy resulted in 1532 hits, from which we retrieved 140 studies after title and abstract review. After full text review, 38 papers were selected to be included in the review (figure 1). The 38 papers we included were published between 2004 and 2015, and the majority (71%) were published between January 2010 and July 2015.

The selected studies came from 15 countries, with five including multiple countries within the same study[31–35]. The majority were from the UK (n=9), the Netherlands (n=7) or Canada (n=5). The studies covered a range of interventions, the most common being drug treatment (n=12), screening (n=8); of which 4 were prenatal screening, provision of services (for example day case surgery[36] or rehabilitation/occupational therapy[37]) and provision of treatment (both for infertility[32,38]). These interventions were based in 26 different indications/diseases, with nine relating to eight different types of cancer[9,31,39–45], three in Down's syndrome[46–48], two relating to kidney disease and organ transplantation[49,50], two in haemophilia[51,52], and two in infertility[32,38] (Table 1).

Table 1: Summary of studies

Disease	**N(%) out of 38 studies**
Cancer	10 (26%)
Chronic Diseases	10 (26%)
Down's Syndrome	3 (8%)
Hemophilia	2 (5%)
Infertility	2 (5%)
Pregnancy and/or Childbirth	2 (5%)
Other (for all <n=2)	9 (24%)
Intervention	
Screening (including prenatal and genetic testing)	9 (24%)
Provision of services (including nondrug treatments – i.e. surgery or occupational therapy)	11 (29%)
Drug Treatment Preferences	16 (42%)
Preferred Symptoms	1 (3%)
Valuation of Health States	1 (3%)
Survey administration	
Self-completed (online)	6 (16%)
Self-completed (postal)	11 (29%)
Self-completed (on location)	6 (16%)
Interview administered	4 (11%)
Mixture of online and postal (one each group)	2 (5%)
Mixture of interview and postal (one each group)	1 (3%)
Mixture of on-location and postal (one each group)	2 (5%)
Mixture of any of the above	7 (18%)
Not reported	None
Number of attributes	**N (%) out of 43 DCEs in 38 studies**

2-4	10 (23%)
5-7	29 (67%)
8-10	4 (9%)
Development of attributes	**N (%) out of 38 studies**
Focus groups	7 (18%)
Interviews or panel consultations (in one or more groups)	15 (40%)
Mixture of focus group and interviews (one method each group)	6 (16%)
Other methods (including mixed methods)	8 (21%)
Not described	2 (5%)
Development of survey	**N(%) out of 38 studies**
Piloting in all groups of respondents	12 (32%)
Piloting in one (but not at all) groups of respondents	4 (11%)
Piloting in neither group of respondents	3 (5%)
Piloted but group not described/unclear	5 (13%)
Not described	14 (37%)

Notes:

1. Studies taking place in multiple countries were listed under each of the countries separately (total % is more than 100%). Multi-country studies took place in Norway and Denmark[33], United States and Canada[31], The Netherlands and Belgium[32], France, Germany, Italy, Spain and the United Kingdom[35], and Japan and the United States[34].

2: One study required participants to complete the DCE twice, once with assistance and once at home soon after[53]; this was entered as both interview and self-complete via post

3: Five studies[9,35,39,54,55] included separate DCEs for the HCP and non-HCP populations; the numbers of attributes for each DCE were entered independently

3.2 Choice and attribute design and piloting

Of the studies we reviewed that reported the process of survey development, 36 (95%) reported the source of attributes used in their DCE and 24 (63%) reported having piloted their study (Table 1 & Appendix 3). The groups that were used to generate attributes and pilot surveys varied. 13 (34%) of the studies that reported their attribute generation sought input from people representative of all groups who would be asked to complete the DCE[31–35,53,54,56–61], and 13 (54%) of the studies that reported the piloting in their study piloted the survey in all respondent groups[31,40,42,45,47,50–52,57,60–63]. There were only five studies that reported having generated their attributes and piloted their survey in all groups of respondents[31,57,60,61,63].

In the 25 studies that did not report having generated attributes using input from all respondent groups, there was an equal split between those that generated attributes using only health care providers (n=7)[39,40,42,50–52,64], non-health care providers (n=9)[36,37,44,46,48,62,63,65,66], or neither (n=7)[9,38,43,45,47,49,67]. Those that reported generating attributes using neither respondent groups most often used literature reviews alone[47,49], or literature reviews in conjunction with expert opinion[38,43,67], information from regulatory requirements[9], or product labelling[45] to inform attributes. Two studies did not report that attributes had been developed in groups representative of the intended respondents; one study reported that attributes and levels were chosen by the authors[55] and the other did not provide any detail[41].

Piloting in all target groups was more common (12 of 24 studies reporting piloting), and a number of studies that did not report generating attributes in all groups reported piloting their survey in all respondents[40,42,47,50,52].

3.3. Attribute classification

There were a total of 230 attributes included in the 38 studies, of which 144 (63%) could be classified as process attributes, 67 (29%) as outcomes, and 19 (8%) as structure. Five studies included two different DCEs and attributes are included from both versions[9,35,39,54,55]. The most common attribute type related to delivery and timing (n=57) and morbidity (n=39), followed by safety (n=29), access (n=26), patient/physician relationships (n=20), mortality (n=16), continuity and co-ordination of care (n=13), health-related quality of life (n=11), legal issues (n=4), infrastructure (n=7), financial issues (n=5), and qualifications of health care providers (n=3). Four DCEs, all looking at issues of screening and testing contained only diagnosis and testing attributes[41,46,48,63], and one looking at varying drug effects in diabetes, contained only morbidity attributes[58].

3.4 Study sample and framing

The DCEs identified tried to compare the preferences of multiple different groups, separating out preferences of health care providers and non-health care providers (Table 2). The composition of the studies was heterogeneous; although just over half of the studies (n=20) compared one group of health care providers with either patients (or in the case of screening, those who had experienced the test)[9,32,34,36,38,40–45,47,54,56,59,61,62,65,66] or the general population[31], 12 studies compared the preferences of patients [33,35,37,46,48,49,51,52,55,57,58] or parents of patients[39] with multiple health care providers, and the remainder varied in either their non-health care provider[53,60,64] or both non-health care provider and health care provider groups[50,63,67].

Table 2: Matrix of preferences sought

	Non-health care		
Health care Professionals	Patients	General public	Parents or caregivers
GP	14 (37%)	5 (13%)	4 (11%)
Dentist	1 (3%)	0 (0%)	0 (0%)
Surgeon	2 (5%)	1 (3%)	2 (5%)
Other physician specialty	12 (32%)	4 (11%)	4 (11%)
Nurse/ nurse specialist	6 (16%)	1 (3%)	3 (8%)
Pharmacist	3 (8%)	0 (0%)	3 (8%)
Other Professions	9 (24%)	2 (5 %)	1 (3%)
Health care trainee	1 (3%)	1 (3%)	1 (3%)

Notes: Where papers sought preferences from multiple stakeholders, these are counted individually. Data show % of times each profession was involved in conjunction with the corresponding non-health care professional group in the 38 studies

The framing of the decision given to the health care providers and non-health care providers was almost equally split between being the same instruction for both groups (n=17)[35,36,38,41–43,45,49–54,57,61,64,67], being different for each group (n=20)[9,31–34,37,39,40,44,46–48,56,58–60,62,63,65,66], or unclear[55].

In studies giving the same instructions to both groups, the question asked respondent to pick between the alternative options provided, but did not provide any specific framing about of who the respondent should assume they were making the decision for[35–38,41–43,45,49,50,53,54,56,57,64,67]. One study did, however indicate that the participants were asked to choose the option with the biggest global benefit, for themselves[52].

Where the instruction was different, the main difference was that the non-health care provider groups made the decision as the patient whilst the health care provider made the decision they would recommend to their patient[9,31–34,39,40,44,46–48,59,60,62,63,65]. One study framed the same vignette about treating a patient in three different ways to different groups, asking regulators which treatment was appropriate, physicians what they would recommend, and patients what they would prefer if they were the patient in the vignette[58]. Two studies took an alternative perspective, asking the health care provider to attempt to predict their patient choice or preference, providing an alternative angle of concordance between patient and health care provider preferences[44,66].

3.5 Analysis: methods, marginal rates of substitution & sub-groups/heterogeneity

The most common analytical methods used were the random-effects probit[36,43,44,46,48,50,51,55,57,65,66], the logit/conditional logit[33,37,47,58,61,63,67], random effects logit[34,40,52,64], multinomial logit[32,35,39,49,53], or mixed logit[9,42,56,60], with five studies using a range of other methods[31,38,45,54,59].

In 23 studies, one or more of the attributes in the DCE was used to scale coefficients of other attributes ; most commonly this was monetary (n=9)[31,36–39,41,52,54,55,59], time (n=9) [33,45–48,50,57,63,67], accuracy of testing (n=4)[47,57,59,63], or risk (n=3)[43,46,56]. One study specified they had framed their cost attribute in a different way for patients (out of pocket payment) and physicians (as additional hospital cost)[64].

The majority of studies (n=34) reported accounting for heterogeneity within samples; this was most commonly analysed using sub-groups[32,33,37,38,44,47,50–53,55,57,59,61,63,64,67], or incorporating respondent information as covariates in the model[9,31,43,60,65]. In other studies, heterogeneity was accounted for by allowing random parameters in the model to be estimated [34,40] or using a heteroskedastic condition logit model[61]. One study reported that heterogeneity in preferences existed because the mean co-efficient for a physician was non-significant, but the standard deviation for the point estimate was significant[42]. Only one study explored different subgroups of respondents using latent class analysis, finding two segments of respondents which differed in their order of preferences but could not be differentiated by their characteristics[54].

3.6. Approaches to measuring concordance

There was no consistent approach to measuring the concordance of preferences between health care providers and non-health care providers. The methods used varied widely, and could be grouped into three broad approaches for descriptive purposes; (1) qualitative comparison of regression coefficients, (2) statistical tests of differences or similarities of coefficients, and (3) regression diagnostics. Despite the variation in methods, all approaches were based on comparisons of concordance at the aggregate level of the sample. One study segmented respondents using latent class analysis[54] but did not explore concordance of patients and health care providers using this approach.

Where coefficients were compared, ranking attributes based on the strength of coefficients[9,35,37–40,42–49,55,61,63,65,66] was the most frequently used method, although there were also approaches based on the difference between coefficients themselves, or the confidence intervals of coefficients[32,51]. An alternative approach to comparing results was to first estimate coefficients on a common scale, using marginal rates of substitution (e.g. using a payment vehicle like willingness to pay[36,38,54], willingness to accept risk[34,56], or time[33]), relative importance[52,64], or some weighting of the model coefficients[60], and then making a comparison of these. One study calculated a ratio of the importance of patient and physician coefficients[34].

In other cases, statistical tests were used to compare the coefficients generated by each of the different groups. These either took the form of correlation-based approaches to look for statistical similarities of coefficients (e.g. Kendall's Tau b[53] or Spearman's rank correlation[43]), unpaired tests for differences in coefficients between groups (e.g. Pearson-chi square[41]), or pooled regression approaches which look for statistical differences between groups based on coefficients[62,65].

Finally, regression based diagnostics were used to assess differences in preferences. These included the use of interaction terms or the Wald test to see if pooled analysis with a respondent group identifier or interaction term was statistically significant[62], or significantly improved model fit[31,34,50,55]. An alternative approach tested whether the coefficients in two regression analyses using different data sets were equal (Chow test) [58]. A third approach was to assess the impact on the scale parameters of different datasets (Swait & Louviere test) [56,57,61] to see whether models would need to be estimated separately if there were underlying differences in the two datasets.

3.7 Findings of concordance

The conclusions from the studies (Table 3) found mixed concordance and discordance most frequently (28 studies), followed by discordance of patient and health care provider preferences (12 studies) and concordant preferences (4 studies). The predominance of mixed concordance and discordance conclusions appear to be consistent irrespective of the methods used to test for concordance.

Table 3: Summary of concordance analysis and resulting conclusions

		Author conclusion		
Method used	N(%)	Evidence of concordance N (%)	Evidence of disagreement N (%)	Mixed (N%)
Qualitative comparison				
Strength of coefficients	19 (50%)	2 (11%)	3 (16%)	14 (74%)
MRS	6 (16%)	-	3 (50%)	3 (50%)
Relative importance	2 (5%)	-	-	2 (100%)
Weighting	1 (3%)	-	-	1 (100%)

Difference	2 (5%)	-	1 (50%)	1 (50%)
Statistical tests				
Similarity	2 (5%)	-	-	2 (100%)
Unpaired differences	1 (3%)	-	-	1 (100%)
Pooled regression	2 (5%)	-	2 (100%)	-
Regression diagnostics				
Wald test/interactions	5 (13%)	-	2 (40%)	3 (60%)
Chow	1 (3%)	1 (100%)	-	-
Swait & Louviere test	3 (8%)	1 (33%)	1 (33%)	1 (33%)

Synthesis of the reported concordance scores from the studies reviewed showed that concordance and discordance varied by the type and classification of attribute (figure 2). The pattern of results indicated that health care professionals appeared to believe that structure and outcome attributes were of greater importance than patients did, whereas patients appeared to place greater importance on process outcomes than health care professionals.

The greatest discordance between patients and health care providers were for mortality (e.g. Chance of dying from lymphoma in 10 years[56]) which health care providers believed to be more important than patients, and infrastructure (e.g. “Number of beds in hospital room”[65]) which health care providers again thought was more important than patients.

Patients placed more importance on issues of process, for example safety (e.g. “Risk of urinary incontinence due to treatment”[42]) than health care providers. Similarly, patients cared more about delivery and timing (e.g. “Route of drug administration”[62]) than health care providers. There was evidence of some discordance around issues of patient and health care provider relationship (e.g. “Physician’s attitude to patients”[32]), morbidity (e.g. “Time necessary to recover (defined as returning to normal activities)”[64]) and access (e.g. Extra cost to patient[55]) with patients rating this as more important than health care providers.

However, within each of these categories there were some attributes which showed evidence of concordance. Within the structure category, there was no evidence of discordance around the qualification of the physician (e.g. “Reputation of surgeon”[43]), within process we found no pattern of discordance between patients and health care providers on aspects of continuity and coordination (e.g.

"Continuity of physicians"[32]), and within outcome no pattern of discordance around health-related quality of life (e.g. Physical quality of life[44]).

4. DISCUSSION

This systematic review found that discordance between patient and health care provider preferences in decisions around health care interventions appears to be more common than concordance, and secondly that concordance (or discordance) is not a binary concept. We identified a large body of work which has attempted to understand the similarity or differences of patient and health care provider preferences. Most of these studies have reported mixed conclusions on the concordance of preferences for patients and health care providers, but there is more evidence of discordance than concordance in the conclusions of these papers. Similarly, we found that concordance or discordance of patient and health care professional preferences, in the DCE context, varies across the different classifications of attributes being considered. In a synthesis of results of the studies we reviewed, our analysis suggested that health care providers place greater importance on attributes of structure and outcomes of care, particularly mortality, than patients do. However, the studies did not provide results which allow us to understand whether and how the importance of different attributes varies within these groups of respondents, which is a limitation of the literature we reviewed.

Another key limitation of the literature we reviewed was that the reasons for differences in preferences between patients and health care providers were unclear. We found considerable variation in the approaches used both between and within the DCEs we reviewed, including methods of analysis and testing for differences or variation in preferences. Almost half of the studies we reviewed used different versions of the survey in patients and health care provider groups, meaning that differing preferences between groups could potentially be attributable to genuine differences, or alternatively to differences in the choice sets they completed. Where versions differed, this was primarily in the perspective respondents were asked to take when indicating their preferences: some were asked to choose from their own perspective, while in others the perspective of patients and perspective of health care providers was different within the same study. For example, patients might be asked to consider their own preferences, while health care providers were asked to try to predict the preferences of their patient. Even in studies that provided the same instructions to both groups, often it was unclear whether the health care provider should be considering their own preferences, the preferences of a patient, or some other preference. Consequently it is unclear whether the results should be expected to be concordant or discordant, and whether the implications of discordant preferences are important.

Only a small number of studies actually provided DCEs with different attributes to different groups of respondents. As we have noted, where DCEs with different framing are used, or the attributes or levels included, comparison of coefficients for the purpose of assessing concordance of preferences is challenging and limited by differences in scales. For this reason, we recommend that any DCE aimed at assessing concordance should use the same attributes and levels, and should report the framing of the question to allow readers to assess whether the tasks are equivalent.

Comparisons of patients and health care provider preferences in the DCEs were also primarily made at the aggregate level, which is not informative about level of agreement and distribution of sub-group preferences. One study used a latent class analysis approach to identify sub-group preferences within patient and health care provider preferences, but did not use this approach to understand whether combinations of patient and health care provider sub-groups had more concordant preferences[54]. As latent class analysis is a relatively new method in the analysis of DCEs, the period covered by our review may predate any increase in published studies applying these methods to understand heterogeneity of preferences within respondent groups. However, there is a need to try to identify groups of patients and health care providers with similar preferences in future DCEs, and opportunities to reanalyze data collected in previously published DCEs to understand preference heterogeneity using these methods. Within groups of patients and health care providers there are likely to be subgroups which are fairly homogenous in their preferences for certain aspects of treatment, but these preferences may be heterogeneous to those of other groups of patients and health care providers. We recommend that studies seeking to understand the concordance of patient and provider preferences should investigate the heterogeneity of preferences within groups and how these relate. Knowing the existence of these groups and their preferences will help determine whether the same treatments, programs or services can be offered to everyone, whether different options can be offered to different groups, or whether patients and health care providers can be more closely matched according to their preferences. Failing to account for heterogeneity in patient and health care provider preferences might mean that a treatment or service could be designed which meets the preferences of the aggregate group, but fails to meet the preferences of sub-groups of people within that population[68]. The implications of our findings are that the health care that people want is often not the same as what health care providers think people want. This lack of concordance suggests that for decisions which involve significant trade-offs (preference-sensitive care), there is a role for eliciting people's preferences and values about their health care options, potentially through tools like decision aids, so health care professionals can offer the most appropriate options[69] or to match health care providers and patients with similar preferences and values.

This study is not without limitations. Firstly, by restricting ourselves to studies using DCE methodology, we took a narrow view of the literature on this topic. However, assessing the consistency of preferences across multiple different methodologies would have introduced considerable additional heterogeneity, making it even more difficult to draw firm conclusions. Secondly, the way we synthesized concordance scores between and across studies required assumptions that are known to be problematic. However, we believe that the value of undertaking a synthesis of results adds to the narrative review of the literature in highlighting the key areas where concordance and discordance is greatest based on the current evidence. For this reason, we believe it is worth making these assumptions, but recognize that this limits the interpretation of this synthesis. Additionally, the terms we used in our literature search strategy meant that, in some cases, papers from studies which set out to compare patient or health care provider preferences, but reported these in separate publications based on a single sample of respondents would not have been incorporated into the review. Finally, it is unclear what the concordance (or lack thereof) of preferences or values means within each type of attribute as regards the patient-centeredness of care or the quality of the treatment decision.

DCE studies provide an excellent opportunity for determining whether there is concordance of values and preferences for aspects of treatments or services between patients and providers. However, our findings highlight that no consistent approach has been taken to understand whether there is concordance, and we have identified a number of issues which have limited the interpretation of the approaches we identified and made recommendations for future studies.

We have also shown that discordance in patient and health care provider preferences appears to be common, and that concordance (or discordance) varies according to which attributes are being considered. For example, for a single decision there could be concordance on the importance of quality of life outcomes, but discordance around mortality outcomes and issues of access. This highlights that concordance should not been considered as a binary outcome and it is important that any measure of value concordance considers all aspects jointly.

Figure 2 Legend

Bar Colours

Solid: concordance score for all attributes classified in this category

Transparent: concordance score for all attributes in sub-category

Concordance Scores

Positive: physician ranks attribute higher than patient

Negative: patient ranks attribute higher than physician

References

1. Stewart, M. Towards a global definition of patient centred care. *BMJ* **322,** 444–445 (2001).
2. Little, P. *et al.* Preferences of patients for patient centred approach to consultation in primary care: observational study. *BMJ* **322,** 468 (2001).
3. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* **268,** 2420–2425 (1992).
4. Haynes, R. B., Devereaux, P. J. & Guyatt, G. H. Physicians' and patients' choices in evidence based practice. *BMJ* **324,** 1350 (2002).
5. Elwyn, G. *et al.* Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* **333,** 417 (2006).
6. Mulley, A. G., Trimble, C. & Elwyn, G. Stop the silent misdiagnosis: patients' preferences matter. *BMJ* **345,** e6572 (2012).
7. Mulley, A., Trimble, C. & Elwyn, G. The problem of silent misdiagnosis. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/patients-preferences-matter-may-2012.pdf. (Accessed: 23rd August 2016)
8. Stacey, D. *et al.* Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst. Rev.* CD001431 (2011). doi:10.1002/14651858.CD001431.pub3
9. Fiebig, D. G., Haas, M., Hossain, I., Street, D. J. & Viney, R. Decisions about Pap tests: What influences women and providers? *Soc. Sci. Med.* **68,** 1766–1774 (2009).
10. Jalukar, V., Funk, G. F., Christensen, A. J., Karnell, L. H. & Moran, P. J. Health states following head and neck cancer treatment: patient, health-care professional, and public perspectives. *Head Neck* **20,** 600–608 (1998).

11. Sassi, F., Mcdaid, D. & Ricciardi, W. Conjoint analysis of preferences for cardiac risk assessment in primary care. *Int. J. Technol. Assess. Health Care* **21,** 211–8 (2005).
12. Casparie, A. F. & van der Waal, M. A. Differences in preferences between diabetic patients and diabetologists regarding quality of care: a matter of continuity and efficiency of care? *Diabet. Med. J. Br. Diabet. Assoc.* **12,** 828–832 (1995).
13. Axel C. Mühlbacher & Christin Juhnke. Patient Preferences Versus Physicians' Judgement: Does it Make a Difference in Healthcare Decision Making? *Patient Prefer. Physicians' Judgement Does It Make Differ. Healthc. Decis. Mak.* **11,** 163–180 (2013).
14. Lancsar, E. & Louviere, J. Conducting discrete choice experiments to inform Healthcare decision making. *Pharmacoeconomics* **26,** 661–677 (2008).
15. Louviere, J. J. & Lancsar, E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Econ. Policy Law* **4,** 527–546 (2009).
16. de Bekker-Grob, E. W., Ryan, M. & Gerard, K. Discrete choice experiments in health economics: a review of the literature. *Health Econ.* **21,** 145–172 (2012).
17. Gold, M. R. *Cost-effectiveness in health and medicine*. (Oxford University Press, 1996).
18. Elliott, R. & Payne, K. *Essentials of Economic Evaluation in Healthcare*. (Pharmaceutical Press, 2005).
19. Reed Johnson, F. *et al.* Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health* **16,** 3–13 (2013).
20. Bridges, J. F. P., Onukwugha, E., Johnson, F. R. & Hauber, A. B. Patient Preference Methods - A Patient Centered Evaluation Paradigm. *ISPOR Connect.* **December, 15,** 4–7 (2007).
21. Mark, T. L. & Swait, J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ.* **13,** 563–573 (2004).

22. Lambooij, M. S. *et al.* Consistency between stated and revealed preferences: a discrete choice experiment and a behavioural experiment on vaccination behaviour compared. *BMC Med. Res. Methodol.* **15,** (2015).
23. Krucien, N., Gafni, A. & Pelletier-Fleury, N. Empirical Testing of the External Validity of a Discrete Choice Experiment to Determine Preferred Treatment Option: The Case of Sleep Apnea. *Health Econ.* **24,** 951–965 (2015).
24. Ryan, M. & Gerard, K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl. Health Econ. Health Policy.* **2,** 55–64 (2003).
25. Harrison, M. *et al.* Risk as an Attribute in Discrete Choice Experiments: A Systematic Review of the Literature. *Patient* **7,** (2014).
26. Donabedian A. The quality of care: How can it be assessed? *JAMA* **260,** 1743–1748 (1988).
27. Mühlbacher, A. C. & Juhnke, C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? *Appl. Health Econ. Health Policy.* **11,** 163–180 (2013).
28. Flynn, T. N., Louviere, J. J., Peters, T. J. & Coast, J. Using discrete choice experiments to understand preferences for quality of life. Variance-scale heterogeneity matters. *Soc. Sci. Med.* **70,** 1957–1965 (2010).
29. Vick, S. & Scott, A. Agency in health care. Examining patients' preferences for attributes of the doctor–patient relationship. *J. Health Econ.* **17,** 587–605 (1998).
30. Malhotra, C., Chan, N., Zhou, J., Dalager, H. B. & Finkelstein, E. Variation in physician recommendations, knowledge and perceived roles regarding provision of end-of-life care. *BMC Palliat. Care* **14,** (2015).
31. Marshall, D. A. *et al.* How do physician assessments of patient preferences for colorectal cancer screening tests differ from actual preferences? A comparison in Canada and the United States using a stated-choice survey. *Health Econ.* **18,** 1420–1439 (2009).

32. Empel, I. W. H. van *et al.* Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. *Hum. Reprod.* **26,** 584–593 (2011).
33. Espelid, I. *et al.* Preferences over dental restorative materials among young patients and dental professionals. *Eur. J. Oral Sci.* **114,** 15–21 (2006).
34. Okumura, K. *et al.* Comparing Patient and Physician Risk Tolerance for Bleeding Events Associated with Anticoagulants in Atrial Fibrillation—evidence from the United States and Japan. *Value Health Reg. Issues* **6,** 65–72 (2015).
35. Chancellor, J., Martin, M., Liedgens, H., Baker, M. G. & Müller-Schwefe, G. H. H. Stated Preferences of Physicians and Chronic Pain Sufferers in the Use of Classic Strong Opioids. *Value Health* **15,** 106–117 (2012).
36. Gidman, W., Elliott, R., Payne, K., Meakin, G. H. & Moore, J. A comparison of parents and pediatric anesthesiologists' preferences for attributes of child daycase surgery: a discrete choice experiment. *Pediatr. Anesth.* **17,** 1043–1052 (2007).
37. Laver, K., Ratcliffe, J., George, S., Lester, L. & Crotty, M. Preferences for rehabilitation service delivery: A comparison of the views of patients, occupational therapists and other rehabilitation clinicians using a discrete choice experiment. *Aust. Occup. Ther. J.* **60,** 93–100 (2013).
38. Huppelschoten, A. G. *et al.* The monetary value of patient-centred care: results from a discrete choice experiment in Dutch fertility care. *Hum. Reprod.* **29,** 1712–1720 (2014).
39. Regier, D. A. *et al.* Discrete Choice Experiment to Evaluate Factors That Influence Preferences for Antibiotic Prophylaxis in Pediatric Oncology. *PLoS ONE* **7,** (2012).
40. Gatta, F. *et al.* Patients' and Physicians' Preferences for Approaches to Bone Metastases Treatment in Turkey. *Int. J. Hematol. Oncol.* **25,** 118–129 (2015).

41. Boone, D. *et al.* Patients' & Healthcare Professionals' Values Regarding True- & False-Positive Diagnosis when Colorectal Cancer Screening by CT Colonography: Discrete Choice Experiment. *PLoS ONE* **8,** e80767 (2013).
42. de Bekker-Grob, E. W. *et al.* Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. *Br. J. Cancer* **109,** 633–640 (2013).
43. Thrumurthy, S. G., Morris, J. J. A., Mughal, M. M. & Ward, J. B. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric cancer. *Br. J. Surg.* **98,** 1124–1131 (2011).
44. Mühlbacher, A. C. & Nübling, M. Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. *Eur. J. Health Econ.* **12,** 193–203 (2011).
45. Shafey, M., Lupichuk, S. M., Do, T., Owen, C. & Stewart, D. A. Preferences of patients and physicians concerning treatment options for relapsed follicular lymphoma: a discrete choice experiment. *Bone Marrow Transplant.* **46,** 962–969 (2011).
46. Bishop, A. J. *et al.* Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. *BJOG Int. J. Obstet. Gynaecol.* **111,** 775–779 (2004).
47. Hill, M., Fisher, J., Chitty, L. S. & Morris, S. Women's and health professionals' preferences for prenatal tests for Down syndrome: a discrete choice experiment to contrast noninvasive prenatal diagnosis with current invasive tests. *Genet. Med.* **14,** 905–913 (2012).
48. Lewis, S. M., Cullinane, F. M., Carlin, J. B. & Halliday, J. L. Women's and health professionals' preferences for prenatal testing for Down syndrome in Australia. *Aust. N. Z. J. Obstet. Gynaecol.* **46,** 205–211 (2006).

49. Davison, S. N., Kromm, S. K. & Currie, G. R. Patient and health professional preferences for organ allocation and procurement, end-of-life care and organization of care for patients with chronic kidney disease using a discrete choice experiment. *Nephrol. Dial. Transplant.* **25,** 2334–2341 (2010).
50. Clark, M. D. *et al.* Who should be prioritized for renal transplantation?: Analysis of key stakeholder preferences using discrete choice experiments. *BMC Nephrol.* **13,** 152 (2012).
51. Mantovani, L. G. *et al.* Differences between patients', physicians' and pharmacists' preferences for treatment products in haemophilia: a discrete choice experiment. *Haemophilia* **11,** 589–597 (2005).
52. Scalone, L., Mantovani, L. G., Borghetti, F., Von Mackensen, S. & Gringeri, A. Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA Study. *Haemophilia* **15,** 473–486 (2009).
53. Bijlenga, D., Birnie, E., Mol, B. W. & Bonsel, G. J. Obstetrical outcome valuations by patients, professionals, and laypersons: differences within and between groups using three valuation methods. *BMC Pregnancy Childbirth* **11,** 93 (2011).
54. Deal, K., Keshavjee, K., Troyan, S., Kyba, R. & Holbrook, A. M. Physician and patient willingness to pay for electronic cardiovascular disease management. *Int. J. Med. Inf.* **83,** 517–528 (2014).
55. Lee, A., Gin, T., Lau, A. S. C. & Ng, F. F. A comparison of patients' and health care professionals' preferences for symptoms during immediate postoperative recovery and the management of postoperative nausea and vomiting. *Anesth. Analg.* **100,** 87–93 (2005).
56. Johnson, F. R. *et al.* Are Gastroenterologists Less Tolerant of Treatment Risks than Patients? Benefit-Risk Preferences in Crohn's Disease Management. *J. Manag. Care Pharm.* **16,** 616–628 (2010).
57. Payne, K. *et al.* Valuing pharmacogenetic testing services: A comparison of patients' and health care professionals' preferences. *Value Health* **14,** 121–134 (2011).
58. Mol, P. G. M. *et al.* Understanding drug preferences, different perspectives. *Br. J. Clin. Pharmacol.* **79,** 978–987 (2015).

59. Beulen, L. *et al.* Women's and healthcare professionals' preferences for prenatal testing: a discrete choice experiment. *Prenat. Diagn.* **35,** 549–557 (2015).
60. Whitty, J. A. *et al.* Preferences for the delivery of community pharmacy services to help manage chronic conditions. *Res. Soc. Adm. Pharm.* **11,** 197–215 (2015).
61. Pedersen, L. B., Kjær, T., Kragstrup, J. & Gyrd-Hansen, D. Do General Practitioners Know Patients' Preferences? An Empirical Study on the Agency Relationship at an Aggregate Level Using a Discrete Choice Experiment. *Value Health* **15,** 514–523 (2012).
62. de Bekker-Grob, E. W., Essink-Bot, M.-L., Meerding, W. J., Koes, B. W. & Steyerberg, E. W. Preferences of GPs and Patients for Preventive Osteoporosis Drug Treatment - Springer. *Pharmacoeconomics* **27,** 211–219 (2009).
63. Hill, M., Suri, R., Nash, E. F., Morris, S. & Chitty, L. S. Preferences for Prenatal Tests for Cystic Fibrosis: A Discrete Choice Experiment to Compare the Views of Adult Patients, Carriers of Cystic Fibrosis and Health Professionals. *J. Clin. Med.* **3,** 176 (2014).
64. Faggioli, G., Scalone, L., Mantovani, L. G., Borghetti, F. & Stella, A. Preferences of Patients, Their Family Caregivers and Vascular Surgeons in the Choice of Abdominal Aortic Aneurysms Treatment Options: The PREFER Study. *Eur. J. Vasc. Endovasc. Surg.* **42,** 26–34 (2011).
65. Neuman, E. & Neuman, S. Agency in health-care: are medical care-givers perfect agents? *Appl. Econ. Lett.* **16,** 1355–1360 (2009).
66. Mühlbacher, A. C., Stoll, M., Mahlich, J. & Nübling, M. Evaluating the concordance of physician judgments and patient preferences on AIDS/HIV therapy - a Discrete Choice Experiment. *Health Econ. Rev.* **3,** 1–11 (2013).
67. Park, M.-H., Jo, C., Bae, E. Y. & 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* **15,** 933–939 (2012).

68. Deal, K. Segmenting Patients and Physicians Using Preferences from Discrete Choice Experiments. *Patient* **7,** 5–21 (2014).

69. Stiggelbout, A. M. *et al.* Shared decision making: really putting patients at the centre of healthcare. *BMJ* **344,** e256 (2012).

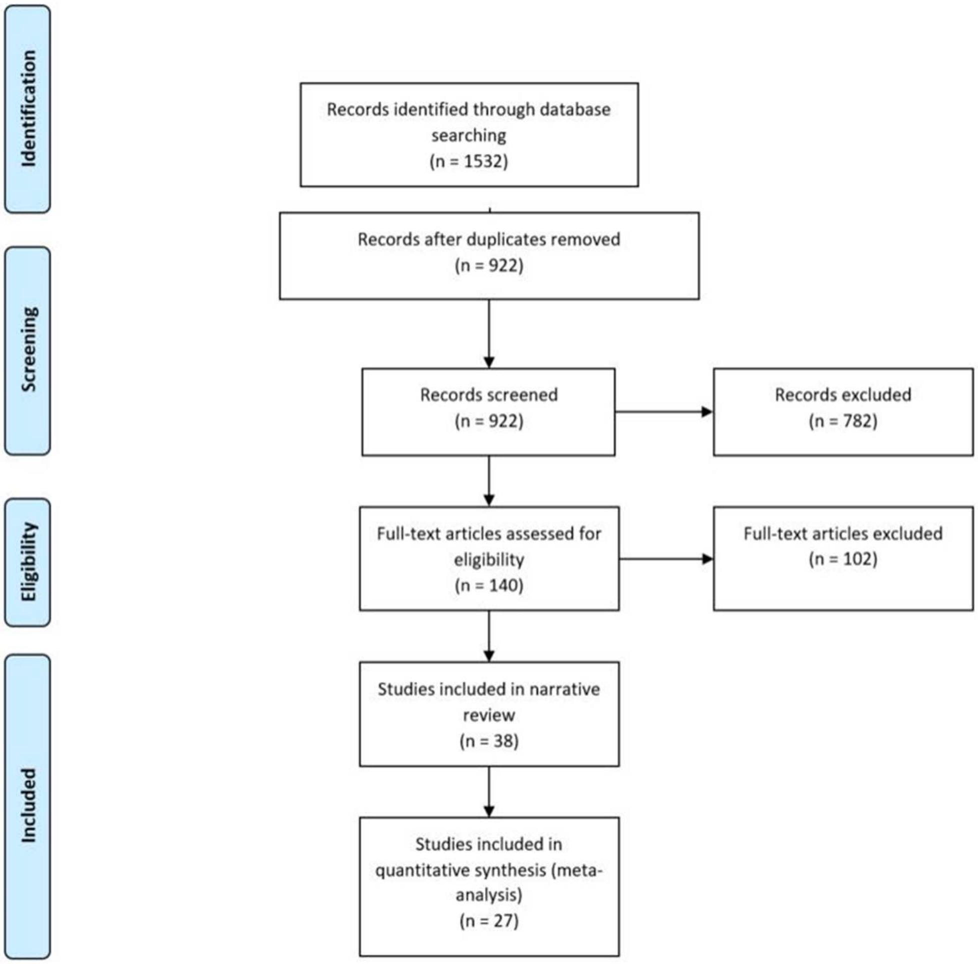


Figure 1: PRISMA Flow Diagram for Systematic Review

211x208mm (96 x 96 DPI)

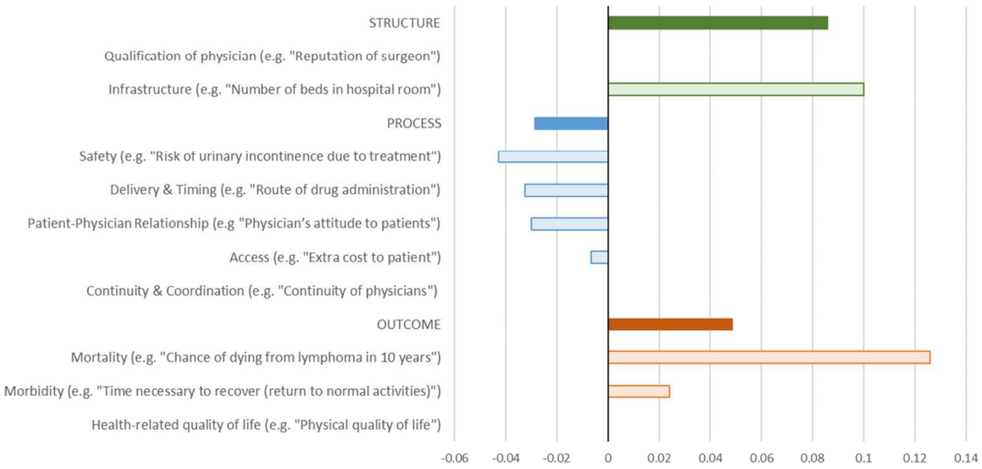


Figure 2: Synthesis of concordance between patients and physician preferences for different types of attributes

366x206mm (96 x 96 DPI)

Appendix 1: Search Terms: MEDLINE July 28th

Patients

patient*.mp OR Patients/ OR women.mp OR men.mp OR parent*.mp OR caregiver.mp OR carer*.mp

Prescribers

prescriber*.mp OR physician*.mp OR provider*.mp OR expert*.mp OR doctor*.mp OR specialist*.mp OR professional*.mp OR surgeon*.mp

Preference

preference*.mp OR value*.mp OR priorit*.mp OR perspective*.mp

Discrete-choice experiment

part-worth utility*.mp OR paired comparison*.mp OR pairwise choice*.mp OR stated preference.mp OR (discrete adj choice$).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] OR (conjoint adj (analys$ or measurement$ or study or studies or choice$)).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

Final Search:

(Patients AND **Preference)** AND **(Prescribers** AND **Preference)** AND **(Discrete-choice experiment)**

Appendix 2: Agreement statistics

		Agreement		
		Overall	Reviewer 2	Reviewer 3
Screening decision		n	n	n
Reviewer 1	Reviewer 2 or 3	922	450	473
Y	Y	68	28	40
Y	N	19	10	9
N	N	820	399	421
N	Y	15	12	3
	N agree	888	427	461
	Agreement %	96%	95%	97%

Appendix 3: Summary of studies

Lead Author	Location	Indication (disease)	Intervention	Samples compared	Attributes Tested	Design of Attributes	Generated in all groups?	Piloting of studies	Piloted in all groups?	Framing of choice tasks	Choice same in all groups?
Bishop et al. 2004	United Kingdom	Down's Syndrome	Screening	253 pregnant women/ 94 HCPs	Time at screening; Detection rate; Risk of miscarriage	Pilot study with 21 women	No	Yes; methods not reported	Not known	Women chose for themselves; HCPs chose for their patients (opt-out and indifference option provided)	No
Lee et al. 2005	China	Postoperative period (1st DCE) and Postoperative nausea/ vomiting(2nd DCE)	Preferred symptoms (1st DCE) and drug treatment (2nd DCE)	200 women undergoing elective surgery/ 52 HCPs	DCE #1: Risk of PONV; Level of Pain; Level of Sedation DCE #2: Type of regimen; Efficacy of antiemetic; Extra cost to patient	Not reported	Not known	Not reported	Not known	DCE #1: Patients chose for themselves; Framing not reported for HCPs. DCE #2: Patients and HCPs randomized into low, moderate, high risk of PONV versions of DCE	No
Manto-vani et al. 2005	Italy	Hemophilia	Drug treatment	178 patients/ 137 HCPs	Perceived viral safety; Risk of inhibitor development; Pharmaceutical dosage form; Distribution mode; Frequency of infusion for prophylaxis; Cost	Generation with physicians, pharmacists and economists; piloting in patients, hematologists and pharmacists. Levels corresponded to available medications.	Yes	Yes: 5 patients, 5 physicians and 5 pharmacists	Yes	Respondents invited to choose one of the two pairs presented	Not reported
Espelid et al. 2006	Norway and Denmark	Dental restoration	Materials used	306 patients/ 107 HCPs	Duration; Appearance; Adverse reaction;	Generated by a general survey of patients and dentists in Great Britain, France, Germany, Italy and Sweden in 1998. Piloted in Norwegian dental students	No	Not reported	Not known	Patients chose for themselves; Dentists chose recommendation for an included patient case; dental assistants chose the best-suited option for the same patient case. Indifference option also.	No
Lewis et al. 2006	Australia	Down's Syndrome	Prenatal screening	113 pregnant women/ 175 HCPs	Timing (weeks); Accuracy (%); Risk (%)	Same method as described in Bishop et al. 2004	No	See Bishop et al. 2004	See Bishop et al. 2004	Patients chose for themselves; HCPs chose what they would offer women	No

Gidman et al. 2007	United Kingdom	Child daycase surgery	Provision of services	280 parents of children undergoing daycase surgery/ 193 HCPs	Parental involvement in medical decisionmaking; Parental presence at the induction of anesthesia; Staff attitude; Postoperative pain; Quality of recovery from anesthesia; Cost to parents	Generated from systematic literature search and analysis of interviews with parents.	No	Yes: in parents of children aged 3-11 years	No	Participants asked to choose the option they thought was preferable.	Yes
de BekkerGrob et al. 2009	The Netherlands	Osteoporosis	Preventative drug treatment	117 patients/ 39 HCPs	Effectiveness of treatment; Nausea as an adverse effect; Total treatment duration; Route of drug administration; Cost	Generated by literature review, expert interviews, and a study in communitydwelling women over 60yo with and without osteoporosis.	Yes	Yes: 2 GPs and 8 patients	Yes	Patients asked to choose for themselves; GPs asked to choose treatments for a standardized female patient over 60yo. 'No-drug' treatment option provided.	No

Fiebig et al. 2009	Australia	Cervical Cancer	Screening	167 women who had Pap tests previously/ 215 HCPs	Women's Survey: Recommended screening interval; Familiarity of GP; Sex of GP*; Time since last cervical screening test; Doctor's recommendations*; Doctor's incentive; Cost of test*; Chance of false negative*; Chance of false positive* GP Survey: Reason for consultation*; Recommended screening interval; Familiarity with patient; Patient's last screening; Age*; Perception of patient's income/ socioeconomic status*; Payment to practice for test	Generated by literature review, current Australian policy context and a pilot test (Fiebig et al 2005).	Reported in Fiebig & Hall 2005	Not reported	Not known	Women asked whether they would choose a cervical Pap and which test; GPs asked whether they would recommend a cervical Pap and which test. Opt-out option provided.	No
Marshall et al. 2009	United States, Canada	Colorectal Cancer	Screening	501 Canadians; 1087 Americans/ 100 HCPs in Canada and the United States	Test process; Test frequency; Requirement for follow-up if initial test is positive; Pain/discomfort from the test; Preparation needed for the test; Risk of complications; Test sensitivity; Test specificity; Cost of test	Generated by a literature review, focus groups and the results of a Canadian-based DCE completed in 2007. Further refined through clinical and methodological input.	Not reported	Yes: patients and physicians attending clinics in California	Yes† †Piloted in American respondents only	Participants asked to choose between two treatments, then between the same two or no treatment. Physicians asked the same, but for a patient aged 50, no history of colorectal cancer	No

Neuman & Neuman 2009	Israel	Labour and Hospitalization following birth	Provision of services	323 women who recently gave birth/ 30 HCPs	Number of beds in hospital room; Attitude of staff towards patient; Professionalism of medical staff; Information given from personnel to patient Travel time from residence to hospital	Generated by literature survey, in-depth interviews with women who recently gave birth, and a pilot study	Not reported	Not reported	Not known	Women asked which maternity ward they would prefer; hospital staff asked to make choices that reproduce and represent the choices made by hospitalized women	Yes
Scalone et al. 2009	Italy	Hemophilia with inhibitors	Drug treatment	37 patients with hemophilia with inhibitors and caregivers/ 64 HCPs	Risk of Infection; Risk of Anamnestic Response; Number of Infusions to stop bleeding; Time to stop bleeding; Time to pain recovery; Number of infusions/week for prophylaxis; Possibility of undergoing major surgery; Increase in healthcare taxes (cost)	Generated by 1 focus group with physicians, pharmacists and health economists; pilot study in 35 patients, pediatric caregivers, physicians and pharmacists; focus groups in physicians to determine levels	Yes	Not reported	Not known	Patients and HCPs asked to choose the option with the maximum global benefit from their point of view	Yes
Davison et al. 2010	Canada	Chronic Kidney Disease (CKD)	Organ procurement, allocation, endof-life care and organisation of care	169 patients with Chronic Kidney Disease / 150 HCPs	Who provides comprehensive, dayto-day care; How deceased donor kidneys should be allocated; How live donor kidneys should be obtained; When should end-oflife care discussions begin; How much information should be provided on prognosis and end-of-life care issues; How should decisions to stop dialysis be made	Generated by review of the literature on aspects of CKD management that are substantial ethical challenges to the nephrology community	No	Not reported	Not known	Participants asked to choose between hypothetical Chronic Kidney Disease programs	Yes

Johnson et al. 2010	United States	Crohn’s disease	Drug treatment	580 patients/ 315 HCPs	Severity of symptoms; Effect on serious complications; Time between flareups; Treatment requires taking oral steroids; Chance of dying from a serious infection within 10 years; Chance of dying or severe disability from PML within 10 years; Chance of dying from lymphoma in 10 years	Generated by review of the literature, consultations with 10 gastroenterologists to finalize hypothetical patient profiles, and interviews with 10 Crohn’s disease patients.	Yes	Yes: 51 Crohn’s disease patients recruited by a market research company.	No	Participants asked which treatment they would choose; Gastroenterologists evaluated for 3 hypothetical patients	No
Bijlenga et al. 2011	The Netherlands	Pregnancy and delivery with complications	Valuation of health states	24 patients+27 laypersons/ 30 HCPs	Maternal health ante partum; Time between diagnosis and delivery; Process of delivery; Maternal outcome; Neonatal outcome	In-depth interviews with 10 patients with gestational diabetes, preeclampsia, and/or intrauterine growth retardation as well as 10 obstetrical care professionals. Attribute levels assigned from these interviews, a literature search and results from the HYPITAT and DIGITAT trials	Yes	Not reported	Not known	Participants asked to choose between pairs of health states	Yes
van Empel et al. 2011	The Netherlands, Belgium	Infertility	Treatment provision	925 patients/ 227 HCPs	Travel time to clinic; Physician’s attitude to patients; Information on treatments; Continuity of physicians; Clinic’s mean ongoing pregnancy rate	Attributes and levels generated from a literature search, focus groups with 82 Belgian and Dutch fertility clinic patients, expert panel of 5 fertility experts.	Yes	Yes: 8 couples during 4 rounds of cognitive interviewing.	No	Patients asked which fertility clinic they would choose. HCPs asked which clinic they would recommend to their patients	No
Faggioli et al. 2011	Italy	Abdominal aortic aneurysm	Drug treatment	160 patients + 102 relatives / 30 HCPs	Type of anesthesia; Time necessary to recover (defined as returning to normal activities);	Attributes and levels generated by review of the literature plus discussion with experienced staff	Yes	Not reported	Not known	Participants asked which treatment would they choose – cost attribute was framed as out-ofpocket cost for	No† †Cost attribute framed differently

					Need to repeat intervention within 5 years; Type of periodical exams and medical visits on follow-up; Risk of severe complications including death; Additional cost	surgeons and health economists with experience in outcomes research. Piloted in 6 patients and 7 experienced staff vascular surgeons				patients and additional hospital cost for HCPs	
Muhlbacher & Nubling 2011	Germany	Multiple Myeloma	Drug treatment	282 patients/ 213 HCPs	Life expectancy/ effectiveness; Adverse effects; Therapy-free intervals; Physical quality of life; Emotional quality of life; Social quality of life; Therapy application; Further treatment options	Reported in Muhlbacher et al. 2008	No	Yes[†] [†]Patient piloting as part of Muhlbacher et al. 2008: also piloted in 30 physicians before this study.	Yes[†] [†]Piloting in patients published in Muhlbacher et al. 2008	Physicians asked to select their patients' preferences; patients' preferences elicited in Muhlbacher et al. 2008	Yes
Payne et al. 2011	United Kingdom	Azathioprine-induced neutropenia	Pharmacogenetic testing services for predicting the risk of azathioprineinduced neutropenia	159 patients prescribed azathioprine/ 138 HCPs	Level of information given; Predictive ability of test; How the sample is collected; Turnaround time for a result; Who explains the test result	Generated by review of the literature; qualitative study described in Fargher et al. using focus groups with HCPs and interviews with patients; expert review of attributes	Yes	Yes: 20 clinic patients and 30 staff	Yes	Participants asked to indicate which test they would choose.	Yes[†] [†]From picture of DCE included
Shafey et al. 2011	Canada	Relapsed follicular lymphoma	Drug treatment	81 patients/ 48 HCPs	Administration of treatment; Toxicity; Remission length; Healthcare cost	Attribute levels were determined by literature review, existing administration protocols and toxicities of each regimen	No	Yes: 2 members of the Calgary Hematology Group, 5 lymphoma patients and 5 medical oncologists	Yes	DCE asked which treatment would participants choose	Yes
Thrumurthy et al 2011	United Kingdom	Esophagogastric cancer	Surgical treatment	81 patients/ 90 HCPs	Mortality; Morbidity; Quality of Life; Cure rate; Hospital type; Reputation of surgeon	Attributes elicited by review of the literature, expert opinion and pilot test	Not known[†] [†]Not explicitly defined whom 'experts' are	Yes: patients	No	Participants chose between two hypothetical surgeries	Yes[†] [†]From picture of DCE included

Chancellor et al. 2012	France, Germany, Italy, Spain, United Kingdom	Chronic Pain	Drug treatment	186 patients/ 310 HCPs	Patient DCE: Effectiveness for pain; Constipation and bowel problems;* Nausea and vomiting;* Alertness;* Energy* Physician DCE: Range of dosage forms;* Proportion of patients with 50% pain reduction; Side effects (constipation, NV, CNS)*	Attributes elicited by review of the literature; focus groups with 44 osteoarthritis/low back pain patients and 40 cancer pain patients; semi structured telephone interviews with 9 physicians.	Yes	Yes: piloted among research colleagues and then soft launched before actual release	Yes	Participants asked to choose between two profiles; optout option provided	Yes
Clark et al. 2012	United Kingdom	Kidney Transplant	Prioritization preferences	908 patients + 41 carers + 48 donors / 113 HCPs	Time spent awaiting transplant; Tissue type matching; Number of child or adult dependents the recipient has; Recipient age; Diseases predominantly affecting life expectancy; Diseases predominantly affecting quality of life	Attributes and level selection mainly informed by discussion with clinicians; piloted in 60 respondents (41 patients, 16 healthcare practitioners, 1 donor, 1 carer, 1 renal consultant's secretary) and analysed using random effects probit	Yes	Yes: 60 respondents (41 patients, 16 healthcare practitioners, 1 donor, 1 carer, 1 renal consultant's secretary); confirmed attributes and survey	Yes	Participants asked to express a stated preference for which one of two transplant recipients should receive a kidney.	Yes
Hill et al. 2012	United Kingdom	Down's Syndrome	Prenatal screening	335 women/ 181 HCPs	Accuracy; Time of results; Risk of miscarriage; Information gained from the test	Attributes selected by literature review	No	Yes: 17 midwives and 20 women	Yes	Patients asked which test they preferred to have; HCPs asked which test they preferred to offer patients. Option to choose neither test provided.	No

Park et al. 2012	South Korea	Metastatic renal cell carcinoma	Drug treatment	140 patients + 60 family members/ 295 HCPs	Progression-free survival; Bone marrow suppression (neutropenia/ thrombocytopenia); Hand-foot skin reaction; GI perforation Bleeding; Administration	Attributes selected by comprehensive literature review and expert opinions.	No	Yes: 20 persons before launch.	Not known	Participants asked which treatment they would choose	No
Pedersen et al. 2012	Denmark	Primary Care	Provision of services	698 members of the general public/ 969 HCPs	Typical waiting time on telephone; Opening hours; Typical waiting time to appointment; Distance to practice; Typical wait time in waiting room; Average consultation time; Who performs routine tasks	Attributes generated by literature review, interviews with GPs, interviews with patients, and discussions with the Organization of General Practitioners of Denmark	Yes	Yes: cognitive interviews and pilot studies in general population and GPs	Yes	Respondents asked to choose their preferred alternative from a set; one set of forced choices and one set of unforced choices	Yes
Regier et al. 2012	Canada	Antimicrobial prophylaxis in pediatric oncology	Provision or nonprovision of drug treatment for prophylaxis	102 parents of. pediatric oncology patients/ 60 HCPs	Risk of infection; Risk of death; Risk of nausea, vomiting, diarrhea or headache; Route of administration; Cost (out of pocket) Two DCE versions: one for antifungal prophylaxis and one for antibacterial	Attributes and levels informed by literature review, qualitative interviews with 3 experienced pediatric oncology physicians	No	Yes: parents and health care professionals	Yes	Parents asked to imagine that their child were a candidate for antimicrobial prophylaxis; HCPs asked to imagine their patients were candidates for prophylaxis. Optout option provided.	Yes
de Bekker-Grob 2013	The Netherlands	Early prostate cancer	Drug treatment	110 patients with + PSA results but no biopsy results yet/ 50 HCPs	Risk of urinary incontinence due to treatment; Risk of erection problems due to treatment; Risk of other permanent side effects due to treatment; Main aim is to cure; Frequency of PSA testing with a risk of new biopsies; Type of treatment	Attributes and levels informed by literature review, interviews with urologists and senior researchers in the field of prostate cancer research.	No	Yes: 11 patients and urologists before launch.	Yes	Participants asked to consider both treatment alternatives as realistic and forced to choose from among them	Yes

Boone et al. 2013	United Kingdom	Colorectal cancer	Screening by CT Colonography	75 patients/ 50 HCPs	Number of additional true positive detections; Number of additional false positive detections	Not reported	Not known	Yes: 10 volunteers	No	Participants asked to choose between a hypothetical "enhanced" test or the standard test	Yes
Laver et al. 2013	Australia	Rehabilitation/ Occupational therapy	Provision of services	100 rehabilitation patients/ 114 HCPs	Mode of therapy; Dose of therapy (per day); Team providing therapy; Amount of recovery made; Cost	Attributes and levels informed by literature search and qualitative interviews with 10 rehabilitation patients.	No	Not reported	Not known	Patients asked to identify their preferred rehabilitation program; HCPs asked to choose what they would recommend for one of their 'typical rehabilitation clients'.	No
Muhlbacher et al. 2013	Germany	HIV/AIDs	Drug treatment	218 patients (from Muhlbacher et al. 2013)/ 131 HCPs	Life expectancy; Long-term side effects; Flexibility of dosing; Physical quality of life; Emotional quality of life; Social quality of life	Attributes and levels informed by literature search and 4 patient focus groups;	No	Yes: 28 patients prior to launch	No	Patients asked to choose between two treatments; HCPs asked to choose how they thought their patients would rate or what they would choose	Yes

Deal et al. 2014	Canada	Cardiovascular disease	Electronic management	74 patients/ 70 HCPs	Patient DCE: Fee/month; Speed of new info added to vascular tracker; Individual patient tracker values displayed; Nurse coordinator tasks /duties; Access to nurse coordinator;* Vascular visits to physician/year Physician DCE: Fee/month; Speed of revised information in the vascular tracker; Tracker values displayed; Nurse coordinator tasks; Nurse coordinator payment/month;* Efficiency in seeing patients;* Billing incentives from government (pay for performance)*	Attributes and levels informed by focus groups conducted with 29 physicians and 21 patients	Yes	Not reported	Not known	Participants considered 18 choice screens including 2 fixed tasks and selected their most preferred out of 3 randomlyselected C3CVT program alternatives	Yes
Hill et al. 2014	United Kingdom	Cystic fibrosis	Prenatal screening	92 adult patients with CF + 50 carriers of CF/ 70 HCPs	Accuracy; Time of results; Miscarriage risk	Attributes and levels informed by a series of focus groups with carriers of single gene disorders.	No	Yes: 20 carriers of cystic fibrosis	No	Patients and carriers chose for themselves; HCPs chose the test they would prefer to offer. Opt-out (neither test) option provided.	No
Huppel-schoten et al. 2014	The Netherlands	Infertility	Treatment provision	550 patients/ 45 HCPs	Clinic's mean ongoing pregnancy rate; Information provision; Patient involvement; Continuity of physicians; Additional costs per IVF cycle	Attributes and levels informed by literature review and an interview with the chief of the healthcare purchasing department in a large Dutch health insurer company.	No	Yes: 13 infertile couples and a health insurer	Yes	Participants asked which clinic they would choose	Yes

Mol et al. 2014	The Netherlands	Diabetes	Drug treatment	226 patients with Type 2 diabetes/ 227 HCPs	Glycated hemoglobin; Cardiovascular disease risk; Effect on body weight; Mild nausea, vomiting or diarrhea; Hypoglycemia; Risk of cancer	Attributes and levels informed by informal literature review, regulatory requirements and product labelling of oral antidiabetic drugs, and 22 indepth interviews with patients, nurses, regulators and pharmacists.	Yes	Not reported	Not known	Standard patient case presented to all participants. Regulators: treatment they felt appropriate. HCPs: treatment they would recommend. Patients: imagine they were that patient to choose	Yes[†] [†]Based on figure of DCE provided
Beulen et al. 2015	The Netherlands	Infant genetic abnormalities	Prenatal screening	507 pregnant women/ 283 HCPs	Minimal gestational age; Time to wait for results; Level of information; Detection rate; False positive rate; Miscarriage risk; Cost	Attributes and levels informed by systematic literature review, semi-structured interviews with pregnant women, and expert panel discussion.	Yes	Yes: 54 participants[†] [†]Type of participant not reported	Not known	Patients asked which test they would prefer to have; healthcare practitioners asked which test they would prefer for their patients. Optout (no test) option provided.	No
Gatta et al 2015	Turkey	Bone metastases	Drug treatment	91 patients/ 99 HCPs	Months to first skeletal-related event/ complication of bone metastases; Months until worsening of pain; Annual risk of Osteonecrosis of the Jaw; Annual risk of renal impairment; Administration regimen	Attributes and levels informed by review of prescribing information, literature review, and consultation with clinical experts	No	Yes: openended interviews with 8 physicians and 15 patients in the United States	Yes	Patients asked to choose the treatment based on their key attributes and the level to which each option fulfilled them; HCPs given two patient cases for this choice	No
Okumura et al. 2015	Japan, United States	Atrial fibrillation	Anticoagulation treatment	Japan: 152 patients/ 164 HCPs United States: 185 patients/107 HCPs	Risk of minor stroke (nondisabling); Risk of major stoke (disabling); Risk of blood clot in the leg (non-CNS, systemic embolism); Risk of heart attack;	Attributes and levels informed by review of clinical trials of anticoagulants, consultation with experts, and semi-structured interviews with 8	Yes	Not reported	Not known	Patients asked to choose between treatments as if choosing for themselves; Physicians asked to choose treatments	No

					Risk of moderate bleeding (clinically relevant, non-major); Risk of nonfatal major bleeding (extracranial major bleed); Risk of all-cause death	patients and 9 physicians in the United States				for 4 virtual patients	
Whitty et al. 2015	Australia	Chronic conditions	Community pharmacy service provision	602 patients or carers / 297 HCPs	Continued medicines supply; Management of ongoing condition; Pharmacy location; Method of getting medicines; Medicine reviews or advice; Average cost per month	Attributes and levels informed by qualitative methods: 97 consumer and carer interviews and 26 focus groups with consumers, carers and health professionals.	Yes	Yes: convenience sample of all target populations, then in 36 adults with chronic conditions before launch.	Yes	Patients and carers asked if they would choose the new service, or remain at their current pharmacy. HCPs asked to choose the pharmacy service they thought consumers would prefer.	No

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary		2 Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6, Appendix
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7, Appendix
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7-8

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7,8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7,8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-11, Tbl. 1-2 Appendix
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-13
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Figure 2
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8-14,16 Appendix 2

Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	11-13, 15 & 16 Figure 2
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	16-19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	17,18
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	17-19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	4

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.