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

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

CrossMark

[1]Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, Canada
[2]Centre for Health Evaluation and Outcome Sciences, St Paul's Hospital, Vancouver, Canada
[3]Department of Medicine, McGill University, Montréal, Canada
[4]Division of Rheumatology, Jewish General Hospital, Montréal, Canada
[5]Lady Davis Institute for Medical Research, Montréal, Canada
[6]School of Population and Public Health, University of British Columbia, Vancouver, Canada

Correspondence to
Dr. Mark Harrison;
mark.harrison@ubc.ca

ABSTRACT

Objectives To review studies eliciting patient and healthcare provider preferences for healthcare interventions using discrete choice experiments (DCEs) to (1) review the methodology to evaluate similarities, differences, rigour 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 healthcare 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 to 31July 2015; discussing a healthcare-related topic; DCE methodology; comparing patients and healthcare providers.

Design Systematic review.

Results We identified 38 papers exploring 16 interventions in 26 diseases/indications. Methods to analyse 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 healthcare provider preferences (n=12) than of concordant preferences (n=4). A synthesis of concordance suggested that healthcare providers rank structure and outcome attributes more highly than patients, while patients rank process attributes more highly than healthcare providers.

Conclusions Discordant patient and healthcare provider preferences for different attributes of healthcare 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 discrete choice experiments (DCEs) 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 focused 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 healthcare provider preferences using DCEs.

► We highlight a lack of consistency within and between studies, which adds difficulty to summarising 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.

BACKGROUND

Patient-centred 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 healthcare providers about their patients' preferences have been termed 'preference diagnosis' and are 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 healthcare providers erroneously deem themselves good at diagnosing their patients' preferences.[7] Discordance between patient and healthcare 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 suggests that patients whose care more closely matches their preferences consume less healthcare.[8] However, evidence is not clear cut on how and to what extent patients and healthcare providers have different preferences. Studies attempting to understand how and how often patient and healthcare provider preferences differ have shown mixed results. For example, some studies suggest no differences between patient and healthcare 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 healthcare provider preferences in healthcare decision-making found that, at the aggregate level, patients and healthcare providers had different preferences.[13] However, the evidence was limited and assessment of the preference discordance remains unclear and complicated due to different methodologies (eg, time trade-off,[10] conjoint analysis,[11] paired comparison), different disease contexts (eg, 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 healthcare 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 healthcare provider values. There is, however, little clarity about how DCEs can be used to assess the concordance of patient and healthcare 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 healthcare provider preferences for healthcare interventions using

DCEs, specifically to (1) review the methodology of DCEs to evaluate similarities, differences and rigour 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 healthcare provider preferences.

METHODS

Systematic search

Search terms were entered into Medline, EMBASE, Econlit, PyscINFO and Web of Science between the dates of 28 July and 31 July 2015. The search strategy combined free text and MeSH terms pertaining to three main concepts-preferences, patients and healthcare 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 online supplementary 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 healthcare 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 (31 July 2015); discussing a healthcare-related topic or condition; eliciting preferences by DCE methodology (modified DCE, rankings, adapted conjoint analysis, conjoint analysis were excluded); and containing a comparison of patients' or their caregivers' preferences to those of healthcare practitioners with prescribing authority using the same DCE questionnaire (ie, 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 online supplementary appendix 2) between all three reviewers, who then met to resolve any disagreement.

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 conceptualising 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 healthcare 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 healthcare practitioners making a diagnosis or recommending a treatment)[26] and outcomes are the effects or consequences of healthcare 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 healthcare framework. Any assumptions made during the process of the classification are declared.

Finally, we extracted details of how patient/non-healthcare practitioner and healthcare 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 was used to inform this assessment of concordance of preferences.

Data synthesis

We attempted to synthesise coefficients derived from each study to observe patterns in attribute types where there was more or less concordance between patients and healthcare 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 healthcare 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 healthcare providers. Since different studies have
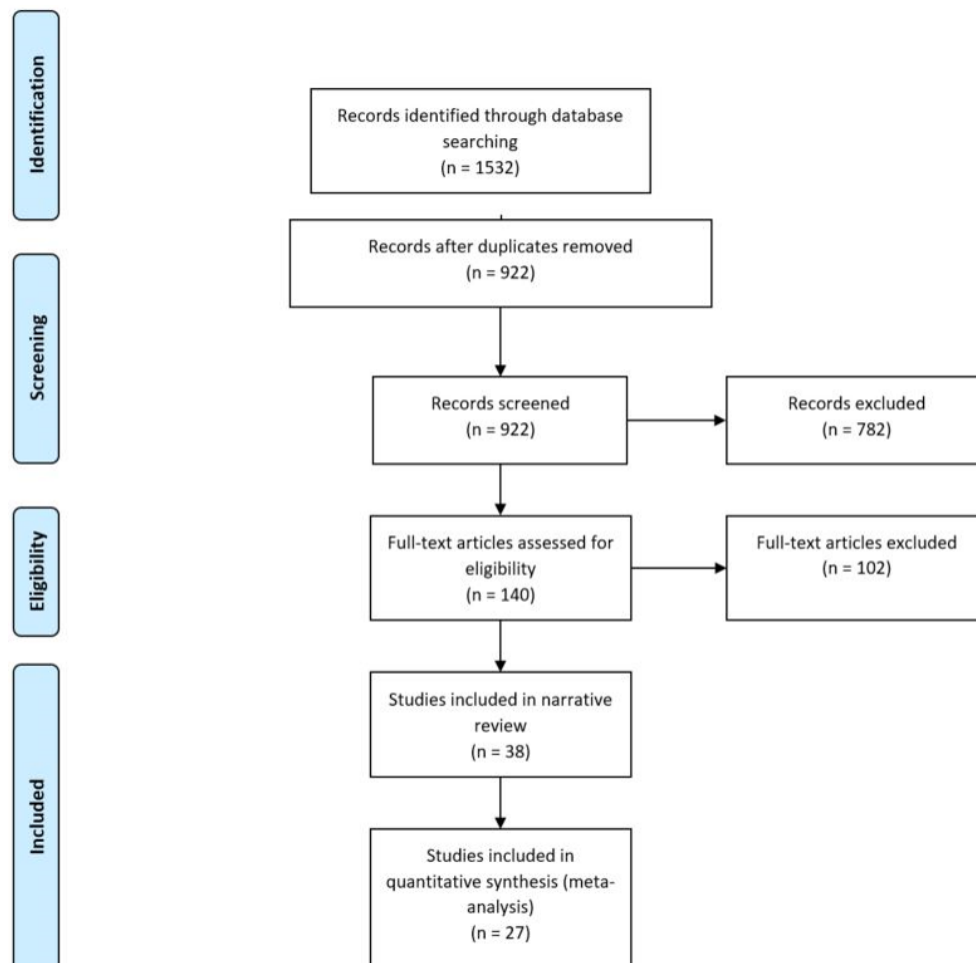

Figure 1 PRISMA diagram for systematic review.

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 healthcare provider believes is most important, as the least important, +1 indicates that a healthcare 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.

RESULTS

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 four were prenatal screening, provision of services (eg, 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 9 relating to 8 different types of cancer,[9 31 39–45] 3 in Down's syndrome,[46–48] 2 relating to kidney disease and organ transplantation,[49 50] 2 in haemophilia[51 52] and 2 in infertility[32 38] (table 1).

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 and online supplementary appendix 3). The groups that were used to generate attributes and pilot surveys varied. Thirteen (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 53–60] 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 56 59–62] There were only five studies that reported having generated their attributes and piloted their survey in all groups of respondents.[31 56 59 60 62]

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 healthcare providers (n=7),[39 40 42 50–52 63] non-healthcare providers (n=9)[36 37 44 46 48 61 62 64 65] or neither

Table 1 Summary of studies

Study Characteristic	Number of studies (%)
Disease	*N (%) out of 38 studies*
Cancer	10 (26%)
Chronic diseases	10 (26%)
Down's syndrome	3 (8%)
Haemophilia	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 non-drug treatments; ie, 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%)

Continued

Table 1 Continued

Study Characteristic	Number of studies (%)
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%)

Studies taking place in multiple countries were listed under each of the countries separately (total % is >100%). Multicountry studies took place in Norway and Denmark,[33] the USA and Canada,[31] the Netherlands and Belgium,[32] France, Germany, Italy, Spain and the UK[35] and Japan and the USA.[34]

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

Five studies[9 35 39 54 67] included separate DCEs for the HCP and non-HCP populations; the numbers of attributes for each DCE were entered independently.

DCE, discrete choice experiment; HCP, healthcare professional.

(n=7).[9 38 43 45 47 49 66] 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 66] 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[67] 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]

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 67] 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 coordination 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 healthcare providers (n=3). Four DCEs, all looking at issues of screening and testing contained only diagnosis and testing attributes,[41 46 48 62] and one looking at varying drug effects in diabetes, contained only morbidity attributes.[57]

Study sample and framing

The DCEs identified tried to compare the preferences of multiple different groups, separating out preferences of healthcare providers and non-healthcare providers (table 2). The composition of the studies was heterogeneous; although just over half of the studies (n=20) compared one group of healthcare 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 55 58 60 61 64 65] or the general population,[31] 12 studies compared the preferences of patients[33 35 37 46 48 49 51 52 56 57 67] or parents of patients[39] with multiple healthcare providers, and the remainder varied in either their non-healthcare provider[53 59 63] or both non-healthcare provider and healthcare provider groups.[50 62 66]

The framing of the decision given to the healthcare providers and non-healthcare providers was almost equally split between being the same instruction for both groups (n=17),[35 36 38 41–43 45 49–54 56 60 63 66] being different for each group (n=20)[9 31–34 37 39 40 44 46–48 55 57–59 61 62 64 65] or unclear.[67]

Table 2 Matrix of preferences sought

Healthcare professionals	Non-healthcare		
	Patients	General public	Parents or caregivers
General practitioner	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%)
Healthcare trainee	1 (3%)	1 (3%)	1 (3%)

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-healthcare professional group in the 38 studies.

In studies giving the same instructions to both groups, the question asked respondents to pick between the alternative options provided, but did not provide any specific framing about or who the respondent should assume they were making the decision for.[35–38 41–43 45 49 50 53–56 63 66] 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-healthcare provider groups made the decision as the patient while the healthcare provider made the decision they would recommend to their patient.[9 31–34 39 40 44 46–48 58 59 61 62 64] 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.[57] Two studies took an alternative perspective, asking the healthcare provider to attempt to predict their patient choice or preference, providing an alternative angle of concordance between patient and healthcare provider preferences.[44 65]

Analysis: methods, marginal rates of substitution and subgroups/heterogeneity

The most common analytical methods used were the random-effects probit,[36 43 44 46 48 50 51 56 64 65 67] the logit/conditional logit,[33 37 47 57 60 62 66] random-effects logit,[34 40 52 63] multinomial logit[32 35 39 49 53] or mixed logit,[9 42 55 59] with five studies using a range of other methods.[31 38 45 54 58]

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 58 67] time (n=9),[33 45–48 50 56 62 66] accuracy of testing (n=4)[47 56 58 62] or risk (n=3).[43 46 55] 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).[63]

The majority of studies (n=34) reported accounting for heterogeneity within samples; this was most commonly analysed using subgroups[32 33 37 38 44 47 50–53 56 58 60 62 63 66 67] or incorporating respondent information as covariates in the model.[9 31 43 59 64] 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.[60] One study reported that heterogeneity in preferences existed because the mean coefficient for a physician was non-significant, but the SD 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]

Approaches to measuring concordance

There was no consistent approach to measuring the concordance of preferences between healthcare providers and non-healthcare 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 healthcare providers using this approach.

Where coefficients were compared, ranking attributes based on the strength of coefficients[9 35 37–40 42–49 60 62 64 65 67] was the most frequently used method, although there were also approaches based on the difference between coefficients themselves or the CIs of coefficients.[32 51] An alternative approach to comparing results was to first estimate coefficients on a common scale, using marginal rates of substitution (eg, using a payment vehicle like willingness to pay,[36 38 54] willingness to accept risk[34 55] or time),[33] relative importance[52 63] or some weighting of the model coefficients,[59] 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 (eg, Kendall's tau b[53] or Spearman's rank correlation),[43] unpaired tests for differences in coefficients between groups (eg, Pearson χ^2)[41] or pooled regression approaches which look for statistical differences between groups based on coefficients.[61 64]

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,[61] or significantly improved model fit.[31 34 50 67] An alternative approach tested whether the coefficients in two regression analyses using different data sets were equal (Chow test).[57] A third approach was to assess the impact on the scale parameters of different data sets (Swait and Louviere test)[55 56 60] to see whether models would need to be estimated separately if there were underlying differences in the two data sets.

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 healthcare provider preferences (12 studies) and concordant preferences (4 studies). The predominance of mixed concordance and discordance conclusions appears to be consistent irrespective of the methods used to test for concordance.

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

Table 3 Summary of concordance analysis and resulting conclusions

Method used	N (%)	Author conclusion		
		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 and Louviere test	3 (8%)	1 (33%)	1 (33%)	1 (33%)

MRS, marginal rate of substitution.

healthcare 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 healthcare professionals.

The greatest discordance between patients and healthcare providers was for mortality (eg, chance of dying from lymphoma in 10 years)[55] which healthcare providers believed to be more important than patients, and infrastructure (eg, 'number of beds in hospital room')[64] which healthcare providers again thought was more important than patients.

Patients placed more importance on issues of process, for example, safety (eg, 'risk of urinary incontinence due to treatment')[42] than healthcare providers. Similarly, patients cared more about delivery and timing (eg, 'route of drug administration')[61] than healthcare providers. There was evidence of some discordance around issues of patient and healthcare provider relationship (eg, 'physician's attitude to patients'),[32] morbidity (eg, 'time necessary to recover (defined as returning to normal activities)')[63] and access (eg, extra cost to patient)[67] with patients rating this as more important than healthcare providers.
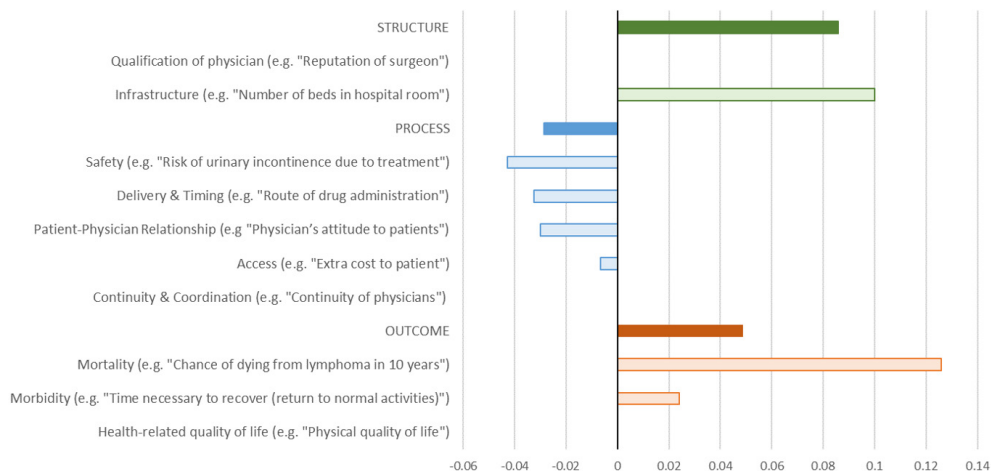

Figure 2 Synthesis of concordance between patient and physician preferences for different types of attributes. Bar colours. Solid: concordance score for all attributes classified in this category. Transparent: concordance score for all attributes in subcategory. Concordance scores. Positive: physician ranks attribute higher than patient. Negative: patient ranks attribute higher than physician.

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 (eg, 'reputation of surgeon'),[43] within process we found no pattern of discordance between patients and healthcare providers on aspects of continuity and coordination (eg, 'continuity of physicians'),[32] and within outcome no pattern of discordance around health-related quality of life (eg, physical quality of life).[44]

DISCUSSION

This systematic review found that discordance between patient and healthcare provider preferences in decisions around healthcare interventions appears to be more common than concordance, and second 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 healthcare provider preferences. Most of these studies have reported mixed conclusions on the concordance of preferences for patients and healthcare 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 healthcare 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 healthcare 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 healthcare 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 healthcare 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 healthcare providers was different within the same study. For example, patients might be asked to consider their own preferences, while healthcare 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 healthcare

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 healthcare provider preferences in the DCEs were also primarily made at the aggregate level, which is not informative about level of agreement and distribution of subgroup preferences. One study used a latent class analysis approach to identify subgroup preferences within patient and healthcare provider preferences, but did not use this approach to understand whether combinations of patient and healthcare provider subgroups 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 healthcare providers with similar preferences in future DCEs, and opportunities to reanalyse data collected in previously published DCEs to understand preference heterogeneity using these methods. Within groups of patients and healthcare 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 healthcare 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, programmes or services can be offered to everyone, whether different options can be offered to different groups or whether patients and healthcare providers can be more closely matched according to their preferences. Failing to account for heterogeneity in patient and healthcare 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 subgroups of people within that population.[68] The implications of our findings are that the healthcare that people want is often not the same as what healthcare 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 healthcare options, potentially through tools like decision aids, so healthcare professionals can offer the most appropriate options[69] or to match healthcare providers and patients with similar preferences and values.

This study is not without limitations. First, 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. Second, the way we synthesised 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 recognise 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 healthcare 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 healthcare 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 be 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* 2001;322:444–5.
2. Little P, Everitt H, Williamson I, *et al*. Preferences of patients for patient centred approach to consultation in primary care: observational study. *BMJ* 2001;322:468.
3. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;268:2420–5.
4. Haynes RB, Devereaux PJ, Guyatt GH. Physicians' and patients' choices in evidence based practice. *BMJ* 2002;324:1350.
5. Elwyn G, O'Connor A, Stacey D, *et al*. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006;333:417.
6. Mulley AG, Trimble C, Elwyn G. Stop the silent misdiagnosis: patients' preferences matter. *BMJ* 2012;345:e6572.
7. Mulley A, Trimble C, Elwyn G. The problem of silent misdiagnosis. http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/patients-preferences-matter-may-2012.pdf (accessed 23 Aug 2016).
8. Stacey D , *et al*. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2011;CD001431.
9. Fiebig DG, Haas M, Hossain I, *et al*. Decisions about pap tests: what influences women and providers? *Soc Sci Med* 2009;68:1766–74.
10. Jalukar V, Funk GF, Christensen AJ, *et al*. Health states following head and neck cancer treatment: patient, health-care professional, and public perspectives. *Head Neck* 1998;20:600–8.
11. Sassi F, McDaid D, Ricciardi W. Conjoint analysis of preferences for cardiac risk assessment in primary care. *Int J Technol Assess Health Care* 2005;21:211–8.
12. Casparie AF, van der Waal MA. Differences in preferences between diabetic patients and diabetologists regarding quality of care: a matter of continuity and efficiency of care? *Diabet Med* 1995;12:828–32.
13. Mühlbacher AC, Juhnke C. 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* 2013;11:163–80.
14. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics* 2008;26:661–77.
15. Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Econ Policy Law* 2009;4:527–46.

16. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21:145–72.

17. Gold MR. *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, Lancsar E, Marshall D, *et al*. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint analysis experimental design good research Practices Task Force. *Value Health* 2013;16:3–13.

20. Bridges JFP, Onukwugha E, Johnson FR, *et al*. Patient preference methods-A patient centered evaluation paradigm. *ISPOR Connec* 2007;15:4–7.

21. Mark TL, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ* 2004;13:563–73.

22. Lambooij MS, Harmsen IA, Veldwijk J, *et al*. Consistency between stated and revealed preferences: a discrete choice experiment and a behavioural experiment on vaccination behaviour compared. *BMC Med Res Methodol* 2015;15:19.

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* 2015;24:951–65.

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* 2003;2:55–64.

25. Harrison M, Rigby D, Vass C, *et al*. Risk as an attribute in discrete choice experiments: a systematic review of the literature. *Patient* 2014;7:151–70.

26. Donabedian A. The quality of care. how can it be assessed? *JAMA* 1988;260:1743CD001431–8.

27. Mühlbacher AC, Juhnke C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? *Appl Health Econ Health Policy* 2013;11:163–80.

28. Flynn TN, Louviere JJ, Peters TJ, *et al*. Using discrete choice experiments to understand preferences for quality of life. Variance-scale heterogeneity matters. *Soc Sci Med* 2010;70:1957–65.

29. Vick S, Scott A. Agency in health care. examining patients' preferences for attributes of the doctor-patient relationship. *J Health Econ* 1998;17:587–605.

30. Malhotra C, Chan N, Zhou J, *et al*. Variation in physician recommendations, knowledge and perceived roles regarding provision of end-of-life care. *BMC Palliat Care* 2015;14:52.

31. Marshall DA, Johnson FR, Kulin NA, *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* 2009;18:1420–39.

32. van Empel IW, Dancet EA, Koolman XH, *et al*. Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. *Hum Reprod* 2011;26:584–93.

33. Espelid I, Cairns J, Askildsen JE, *et al*. Preferences over dental restorative materials among young patients and dental professionals. *Eur J Oral Sci* 2006;114:15–21.

34. Okumura K, Inoue H, Yasaka M, *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* 2015;6:65–72.

35. Chancellor J, Martin M, Liedgens H, *et al*. Stated preferences of physicians and chronic pain sufferers in the use of classic strong opioids. *Value Health* 2012;15:106–17.

36. Gidman W, Elliott R, Payne K, *et al*. A comparison of parents and pediatric anesthesiologists' preferences for attributes of child daycase surgery: a discrete choice experiment. *Paediatr Anaesth* 2007;17:1043–52.

37. Laver K, Ratcliffe J, George S, *et al*. 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* 2013;60:93–100.

38. Huppelschoten AG, Verkerk EW, Appleby J, *et al*. The monetary value of patient-centred care: results from a discrete choice experiment in dutch fertility care. *Hum Reprod* 2014;29:1712–20.

39. Regier DA, Diorio C, Ethier MC, *et al*. Discrete choice experiment to evaluate factors that influence preferences for antibiotic prophylaxis in pediatric oncology. *PLoS One* 2012;7:e47470.

40. Gatta F, *et al*. Patients' and Physicians' Preferences for Approaches to Bone Metastases Treatment in Turkey. *Int J Hematol Oncol* 2015;25:118–29.

41. Boone D, Mallett S, Zhu S, *et al*. Patients' & healthcare professionals' values regarding true- & false-positive diagnosis when colorectal Cancer screening by CT colonography: discrete choice experiment. *PLoS One* 2013;8:e80767.

42. de Bekker-Grob EW, Bliemer MC, Donkers B, *et al*. Patients' and urologists' preferences for prostate Cancer treatment: a discrete choice experiment. *Br J Cancer* 2013;109:633–40.

43. Thrumurthy SG, Morris JJ, Mughal MM, *et al*. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric Cancer. *Br J Surg* 2011;98:1124–31.

44. Mühlbacher AC, 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* 2011;12:193–203.

45. Shafey M, Lupichuk SM, Do T, *et al*. Preferences of patients and physicians concerning treatment options for relapsed follicular lymphoma: a discrete choice experiment. *Bone Marrow Transplant* 2011;46:962–9.

46. Bishop AJ, Marteau TM, Armstrong D, *et al*. Women and health care professionals' preferences for down's Syndrome screening tests: a conjoint analysis study. *BJOG* 2004;111:775–9.

47. Hill M, Fisher J, Chitty LS, *et al*. 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* 2012;14:905–13.

48. Lewis SM, Cullinane FM, Carlin JB, *et al*. Women's and health professionals' preferences for prenatal testing for Down syndrome in Australia. *Aust N Z J Obstet Gynaecol* 2006;46:205–11.

49. Davison SN, Kromm SK, Currie GR. 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* 2010;25:2334–41.

50. Clark MD, Leech D, Gumber A, *et al*. Who should be prioritized for renal transplantation?: analysis of key stakeholder preferences using discrete choice experiments. *BMC Nephrol* 2012;13:152.

51. Mantovani LG, Monzini MS, Mannucci PM, *et al*. Differences between patients', physicians' and pharmacists' preferences for treatment products in Haemophilia: a discrete choice experiment. *Haemophilia* 2005;11:589–97.

52. Scalone L, Mantovani LG, Borghetti F, *et al*. Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA Study. *Haemophilia* 2009;15:473–86.

53. Bijlenga D, Birnie E, Mol BW, *et al*. Obstetrical outcome valuations by patients, professionals, and laypersons: differences within and between groups using three valuation methods. *BMC Pregnancy Childbirth* 2011;11:93.

54. Deal K, Keshavjee K, Troyan S, *et al*. Physician and patient willingness to pay for electronic cardiovascular disease management. *Int J Med Inform* 2014;83:517–28.

55. Johnson FR, Hauber B, Özdemir S, *et al*. Are gastroenterologists less tolerant of treatment risks than patients? Benefit-risk preferences in Crohn's disease management. *J Manag Care Pharm* 2010;16:616–28.

56. Payne K, Fargher EA, Roberts SA, *et al*. Valuing pharmacogenetic testing services: a comparison of patients' and health care professionals' preferences. *Value Health* 2011;14:121–34.

57. Mol PG, Arnardottir AH, Straus SM, *et al*. Understanding drug preferences, different perspectives. *Br J Clin Pharmacol* 2015;79:978–87.

58. Beulen L, Grutters JP, Faas BH, *et al*. Women's and healthcare professionals' preferences for prenatal testing: a discrete choice experiment. *Prenat Diagn* 2015;35:549–57.

59. Whitty JA, Kendall E, Sav A, *et al*. Preferences for the delivery of community pharmacy services to help manage chronic conditions. *Res Soc Adm Pharm* 2015;11:197–215.

60. Pedersen LB, Kjær T, Kragstrup J, *et al*. Do general practitioners know patients' preferences? an empirical study on the agency relationship at an aggregate level using a discrete choice experiment. *Value Health* 2012;15:514–23.

61. de Bekker-Grob EW, Essink-Bot M-L, Meerding WJ, *et al*. Preferences of GPs and patients for Preventive Osteoporosis Drug Treatment. *Pharmacoeconomics* 2009;27:211–9.

62. Hill M, Suri R, Nash EF, *et al*. 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* 2014;3:176–90.

63. Faggioli G, Scalone L, Mantovani LG, *et al*. 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* 2011;42:26–34.

64. Neuman E, Neuman S. Agency in health-care: are medical caregivers perfect agents? *Appl Econ Lett* 2009;16:1355–60.

65. Mühlbacher AC, Stoll M, Mahlich J, *et al*. Evaluating the concordance of physician judgments and patient preferences on AIDS/HIV therapy - a Discrete Choice Experiment. *Health Econ Rev* 2013;3:30–11.

66. Park MH, Jo C, Bae EY, *et al*. A comparison of preferences of targeted therapy for metastatic renal cell carcinoma between the patient group and health care professional group in South Korea. *Value Health* 2012;15:933–9.

67. Lee A, Gin T, Lau AS, *et al*. 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* 2005;100:87–93.

68. Deal K. Segmenting patients and physicians using preferences from discrete choice experiments. *Patient* 2014;7:5–21.

69. Stiggelbout AM, Van der Weijden T, De Wit MP, *et al*. Shared decision making: really putting patients at the centre of healthcare. *BMJ* 2012;344:e256.