Michela Tinelli

Applying discrete social experiments in social care research

Monograph (Other)

Original citation:
Tinelli, Michela (2016) *Applying discrete social experiments in social care research*. Method Reviews, 19. NIHR, SSCR, London, UK.

Originally available from School for Social Care Research

This version available at: http://eprints.lse.ac.uk/73723/

Available in LSE Research Online: April 2017

School for
Social Care Research

NHS
National Institute for
Health Research

APPLYING DISCRETE CHOICE EXPERIMENTS IN SOCIAL CARE RESEARCH

METHODS REVIEW 19

DR MICHELA TINELLI

LSE Enterprise

The School for Social Care Research

The School for Social Care Research was set up by the National Institute for Health Research (NIHR) to develop and improve the evidence base for adult social care practice in England in 2009. It conducts and commissions high-quality research.

About the author

Dr Michela Tinelli is Assistant Professorial Research Fellow within the Personal Social Services Research Unit at the London School of Economics and Political Science. She was a researcher within LSE Enterprise at the time of writing this review.

Michela's research is primarily concerned with the evaluation of patient experience, burden of the disease and the socio economic impact of healthcare interventions in individuals with chronic illnesses and hard to serve populations (i.e. homeless people). She is working on an NIHR-funded study comparing hospital discharge arrangements for homeless people in England, and on a European Brain Council funded project looking at the socio-economic impact of coordinated and patient-centered clinical interventions in brain disorders in Europe. She is also working at LSE Enterprise on an international project comparing the burden of the disease and early management in multiple sclerosis across healthcare settings. A quantitative researcher, she has specialist skills in stated preferences techniques and other benefit evaluations in health policy developments, including clinical outcomes, health related quality of life, and patient satisfaction, which she has applied in benefit and economic evaluations. She has also expertise in survey development and design.

NIHR School for Social Care Research
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE

Email: sscr@lse.ac.uk

Tel: +44 (0)20 7955 6238

Website: www.sscr.nihr.ac.uk

This report presents an independent review commissioned by the Department of Health's NIHR School for Social Care Research. The views expressed in this publication are those of the authors and not necessarily those of the NIHR School for Social Care Research or the Department of Health, NIHR or NHS.

Applying discrete choice experiments (DCEs) in social care research

ABSTRACT

Discrete choice experiments (DCEs) have been widely used by economists to elicit people's values in a number of areas, including market, transport and environmental issues. The last two decades have seen an increasing use of the technique in health economics, and it is beginning to be applied in social care and related research. This review aims to help social care researchers, policymakers and practitioners make the best use of DCEs to value preferences in social care settings. It discusses what DCE is, what you can do with it, and its use to incorporate informal care in economic evaluations. It also describes the key stages of developing a DCE for social care and presents a comprehensive search of the literature to identify and describe DCE applications to social care. Some of the important challenges of applying DCEs to social care are identified, and the need for further methodological development is discussed.

RECOMMENDATIONS

- DCE can be used within social care as a useful tool to inform the evaluation of social care interventions and social-care related quality of life and wellbeing measures.
- Different methods of data collection (such as face-to-face interviews) in different settings and/or with different client groups should be explored as potential means to improve the utility of the approach in this setting.
- The benefit of using pictorial representations of choice sets (to reduce cognitive burden) should be explored further with respondents.
- Careful attention should be placed on the determination of choice sets and selection and levels of attributes within DCE design to reduce the potential of bias in the estimates produced.
- The cognitive burden to respondents should be reduced when using DCE within social care to help respondents choose options that reflect their true preferences.
- Using conditional logit modelling for analysis could improve insights about variation of responses with clear relevance for policymakers.
- Further research is needed to test the external validity of DCE responses in a social care setting which can improve and inform best practice guidance.
- Further comparative research of DCE with best-worst scaling (an alternative stated preference approach) is needed to determine which is the superior approach.

KEYWORDS

discrete choice experiment, stated preferences, social care research

ACKNOWLEDGEMENTS

The author is grateful to Mike Clark and Martin Knapp within NIHR SSCR for their edits on earlier drafts of this review, and for the external reviewers for their comments.

Contents

Introduction

An economic evaluation in social care should adopt a broad societal perspective to consider the impact of an intervention on all relevant stakeholders, including not only the provider(s) but also people who use services and their carers (Drummond *et al.* 2005, NICE 2013, SCIE *et al.* 2011, van den Berg *et al.* 2005) and other societal impacts.

In the context of this paper social care refers to the system of care and support for a range of practical help that people need arising from, for example, an illness or disability, rather than to the immediate health care needs arising from a condition. In England this may be funded by the state or, increasingly, from people's own finances (referred to as self-funded care). Support may be delivered in communities, people's homes, or in institutional care, and can be delivered by people working in the statutory, independent (for profit) or third (charity) sectors of the economy. Unpaid carers, usually relatives of the person needing support but not always, are a key component of this care system.

Service users and their carers, then, are at the centre of the social care system, and it is therefore critical that evaluation and planning of social care services should include their costs and that the outcomes of interventions should be valued from their perspective. Different approaches have been employed to value social care and carer support and have been extensively described and discussed elsewhere (Faria *et al.* 2012, Flynn *et al.* 2007, Koopmanschap *et al.* 2008, Marley *et al.* 2008, Netten 2011, Oremus and Tarride 2008, van den Berg *et al.* 2008, 2005, 2004).

The focus of this review is on the discrete choice experiment (DCE) approach, an economics tool that has been widely used to elicit people's values in a number of areas, including market, transport and environmental issues (Louviere *et al.* 2010). The last two decades have seen an increasing use of the technique in health economics, and, more recently, its application to social care and related research (de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003). A few examples include its use to value carer tasks (Mentzakis *et al.* 2011), support services to aid the development of end-of-life health care services (Hall *et al.* 2013) and long-term care services (Nieboer *et al.* 2010), and to inform the development of preference-weighted outcome measures, for example the Adult Social Care Outcome Toolkit (ASCOT) – see www.pssru.ac.uk/ascot/.

The overall aim of this paper is to provide social care researchers, policymakers and practitioners with best practice guidelines about how to use DCE to value people's preferences in social care settings. The objectives are: describing DCE and its use to value benefits in the delivery of social care; providing guidance on the key stages of developing a DCE for social care; undertaking a comprehensive search of the literature to identify and describe DCE practice in social care research; and providing recommendations for the development of DCE to support its application to social care research.

The next section presents what DCE is and, as an example of its application, what it is possible to do with it to incorporate carer support in economic evaluations. Following that, the paper covers the different phases of a DCE and application to social care. Results from a literature review of DCE use in social care research are presented and discussed in the light of the DCE evidence from healthcare. The last section reports recommendations for further development of DCE to support its application to social care.

DCE and measuring benefits in the delivery of social care

The discrete choice experiment (DCE) approach seeks to establish the relative importance to people of different characteristics in the provision of a good or service (de Bekker-Grob *et al.* 2012). It assumes that any good/service can be defined as a combination of levels of a given set of attributes and the total utility (satisfaction or preference) that an individual derives from that product is determined by those attributes.

In the example shown in figure 1, the choice of renting alternative accommodation to that currently occupied may depend on different factors, including type of accommodation (flat, semi-detached house, detached house), number of rooms, facilities available in the accommodation (such as gas/electric heating), closeness to work (miles), facilities available nearby (shops, schools, pubs), availability for moving in (weeks), and monthly rent (£).

The outcomes of a DCE provide an alternative measure of benefit, where weights can be attached to a range of outcomes (e.g. the level of pain accepted during rehabilitation in a healthcare study), and to process attributes (e.g. location of care, frequency of care) (de Bekker-Grob *et al.* 2012). For example a DCE provides opportunities to evaluate whether or not a given additional services and/or cash benefits are important to caregivers to support their caregiving role; the relative importance of these various attributes; and the trade-offs individuals are willing to make between them. It presents respondents with a set of choices between hypothetical scenarios (i.e. combinations of attributes) picked from all possible choices determined according to statistical design properties. Each choice includes two or more alternatives which vary in the levels of the attributes of interest, and individuals are asked to choose one alternative (see example in Figure 1). Usually each respondent is asked to make a number of choices within a single survey questionnaire.

FIGURE 1 : Example of a DCE choice question (from Tinelli, 2007)

	PROPERTY A	PROPERTY B	
Type of property	Flat	Semi-detached house	
Number of rooms	2	4	
Fully furnished	Yes	No	
Closeness to work (miles)	1	2	
Facilities available nearby	Only shops	Shops, schools and bars	
Monthly rent (£)	450	750	
Please place a tick on the property you would choose	Property A ☑	Property B ☐	Current* ☐

* Information on current property to be collected from each individual within the DCE questionnaire

A DCE's usefulness is that it will identify what characteristics of the service respondents value. For example, in Figure 1, they could value 'rent', 'closeness to work', 'whether it is already furbished', 'number of rooms', 'facilities available nearby' rather than 'type of accommodation' and 'availability for moving in'. It can also identify the relative values that they attach to these characteristics (for example, they could value 'rent' and 'closeness to work' more than other aspects) and the trade-offs they are willing to make (e.g. how much extra time they are willing to wait in order to rent a place within one mile of work).

DCEs allow estimation of a utility (or satisfaction) function which specifies the relationship between the service attributes and consumer preferences. From this function it is possible to estimate utility or satisfaction scores for alternative accommodation on offer. Such utility scores can be combined with costing data to make recommendations concerning the most 'efficient' way of providing the good/service (Drummond *et al.* 2005).

When a cost is included, we can calculate how much respondents are willing to pay (WTP) for changes in attributes levels (for example, how much they are willing to pay to rent a flat closer to work), or for an overall change in the service (how much they are willing to pay to move out from their present property (with its specific attributes levels) to an alternative that is, for example, a flat with three rooms, fully furnished, two miles from work, with shops the only facility nearby, and available in three weeks' time at a rent of £750).

Other examples, this time in the context of social care, include the use of DCEs to develop a social care outcome tool able to capture information about an individual's social care-related quality of life (SCRQoL) (Netten *et al.* 2012), and to value population preferences for long-term care services (Nieboer *et al.* 2010). The former provides the basis for a social care equivalent of the influential quality-adjusted life-year (QALY) measure used in economic evaluation in health care to assess and compare the value for money of different interventions (see the ASCOT toolkit (Netten *et al.* 2012)). The eight domains of the SCRQoL include: 'basic' aspects of quality of life (personal cleanliness and comfort, accommodation cleanliness and comfort, food and drink, and feeling safe); the impact of the way the care and support are provided on people's self-esteem (dignity); and social participation, occupation and control.

The latter example describes population preferences for long-term care. The DCE respondents were asked to choose the better of two care scenarios for four groups of hypothetical individuals who use social care services: frail and demented elderly people, with and without partners (see Box 1). The scenarios described long-term care using the following ten attributes: hours of care, organised social activities, transportation, living situation, same person delivering care, room for individual preferences, coordination of services, punctuality, time on waiting list, and co-payments (see Figure 2).

BOX 1 : Different profiles of people who use social care services (Nieboer et al. 2010)

Try to picture that you are acquainted with an elderly person suffering from various physical problems. Outside of the house he or she cannot walk independently (not even with walking aids), and already has experienced several falls. Medications are taken for a number of health complaints. Getting dressed independently is still possible but takes a lot of time and energy. This person lives alone and has contact with only a few people. Although undertaking things alone is out of the question, the elderly person likes to keep active and mean something to others. To help him or her to achieve this goal, professional help is being sought.

Try to picture that you are acquainted with an elderly person suffering from various physical problems. Outside of the house he or she cannot walk independently (not even with aids), and has already experienced several falls. Medications are taken for a number of health complaints. Getting dressed independently is still possible but it takes a lot of time and energy. This person is married, but other than that has contact with only a few people. Although undertaking things alone is out of the question, the elderly person likes to keep active and mean something to others. To help him or her to achieve this goal, professional help is being sought.

Try to picture that you are acquainted with someone who lives with dementia. This person lives alone and tends to forget things that just happened. For example, a phone call or someone dropped by in the morning. Daily activities such as making coffee and doing the dishes are still manageable. But gradually he or she becomes more dependent on others for the daily activities. As far as possible neighbours and friends take care, but this is a heavy task also because it requires constant supervision. The dementia patient is depressed and often reacts in a snappy and angry manner. To help this person professional help is being sought.

Try to picture that you are acquainted with someone who lives dementia. This person is married and tends to forget things that just happened. For example, a phone call or someone dropped by in the morning. Daily activities such as making coffee and doing the dishes are still manageable. But gradually he or she becomes more dependent on others for the daily activities. As far as possible the partner takes care, but this is a heavy task also because it requires constant supervision. The person with dementia is depressed and often reacts in snappy and angry manner. To help the couple, professional help is being sought.

FIGURE 2 : Example of DCE application to social care (Nieboer et al. 2010)

Please indicate which care package you find most suitable for the described person*. You will find a description of the content of care and modes of service delivery when you click on the aspect in question.

	PACKAGE A	PACKAGE B
Hours of care per week	2 hours	12 hours
Participation in organised social activities	Not available	One half day per week
Transportation services	Available	Not available
Living situation	Living independently at home	Sheltered accommodation
Person delivering care	Varying care providers	Regular care providers
Content of care	Standardised care	The content of care is determined individually
Coordination of services.	Have to arrange little	Have to arrange a lot.
Punctuality	Maximum weight 15 minutes	Maximum weight 1 hour
Time on waiting list	12 months	8 months
Co-payments per week	None	€ 50
Most suitable care package is:	☐	☐

*See Box 1 : Person profiles for DCE experiment.

Conducting a DCE exercise

Identifying the decision to be made is the most important stage of the study. The researchers seek to characterise the decision problem in terms that the decision-maker understands. The researchers need to: define the choice situation (e.g. care services available to people with long-term care needs); search for information on alternatives (e.g. care packages), attributes (e.g. transportation services) and their levels (e.g. available, not available); and construct choice sets (see figure 2). These items are crucial to model the stages of decision-making so that they closely match those that individuals make in real life. When the choice being studied is less familiar to the respondent, proper characterisation of the decision is crucial in order to maximise the chances of communicating the desired information to the respondents (for example, see box 1). The researchers also need to identify sources of individual heterogeneity (e.g. age, sex, income, rural/urban area, health state) that could lead to important differences in values, attitudes, beliefs, and preferences.

The design of the DCE is crucial as it influences how much information can be extracted from respondents' preferences. The construction of the options to choose from, or the choice set (see example in figure 2) must be based on the main elements that influence the choice modelled. These elements can be divided into four stages: (1) establishing attributes and their levels; (2) creating a choice set and measuring design efficiency (efficiency indicates the ability of the design in capturing information, see more below); (3) constructing the survey instrument and collecting data; (4) analysing the DCE data. Guidance on how to construct an experimental design is available from the health economic literature (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012, Lancsar and Louviere 2008, Ryan and Gerard 2003), and from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidance on Good Research Practices for conjoint analysis (Bridges *et al.* 2011, Reed Johnson *et al.* 2013). Establishing the attributes and their levels

Deriving attributes and levels is one of the most important stages when conducting a DCE. The most efficient experimental design and advanced modelling of analysis cannot compensate if the attributes and levels are not appropriate. At this stage the researcher needs to understand the different aspects covered by the research topic in order to accurately define the spectrum of attributes and levels describing the experiment. These need to be measurable and meaningful for the respondent. A misspecification of the attributes and attribute-levels has negative implications for the design and implementation of DCEs and creates a risk of producing biased DCE results, which can mislead interpretation of findings for policy implementation and practice changes.

For example, the inclusion of the cost attribute plays an important role in the experiment as it allows the estimation of benefits (satisfaction, utility) derived from a particular service in monetary terms (£) – the willingness to pay (WTP). Setting appropriate levels (both in number and range) for the cost attribute is an important aspect in the estimation of WTP. There are alternative ways of presenting the cost attribute that are often used: the definition of the cost attribute as a discount (willingness to accept) or as a payment (willingness to pay) (Grutters *et al.* 2008). Monetary values of attributes can also be indirectly estimated by including waiting time as an attribute. Ryan *et al.* (2001) estimated the value of attributes in terms of time, and this was converted to a monetary measure using the value of waiting time for public transport.

In general, the levels chosen for the attributes must be plausible and acceptable, and must allow for trade-offs between combinations of the attributes. For example, a respondent might prefer to rent a flat close to work and available immediately, but the lower rent associated with an alternative option might make him/her choose differently. A review of the literature by Marshall and colleagues (2010) estimated that 70% of the studies used three to seven attributes, with most studies having six attributes, or four attribute levels.

This possible restriction on the number of attributes refers to an assumption of the DCE approach that individuals consider all the attributes, and make trade-offs (e.g. how much money an individual would be willing to give up to have ‘facilities available nearby’). The number of feasible choice sets will also restrict the number of attributes and levels if the full survey is used in a single sample. Previous studies have demonstrated that respondents are able to manage up to 16 choices (de Bekker-Grob *et al.* 2012).

There are alternative approaches to be applied when framing the choice set if it is important to use a higher number of attributes and levels than is feasible in a single survey. When there is a relatively large number of attributes one option reported in the social care literature is to split them across two or more separate DCE sets to allow the full set of domains to be examined (Burge *et al.* 2010, Hall *et al.* 2013). Rather than conducting this as two separate studies, each looking at one of the subset of domains, Netten *et al.* (2012) set up a single experiment in which each respondent was asked to consider both subsets, with the associated advantage that the samples for the two sub-studies were matched. Another possible approach to reduce the number of choices to a manageable number would involve splitting subsets of choices into separate surveys (see the blocking approach below).

Attribute development has to be rigorous, systematic, and transparently reported (Coast *et al.* 2012). Various methods have been applied to the development of DCE attributes. These include literature reviews, existing conceptual and policy-relevant outcome measures, theoretical arguments, panel of experts, patient surveys, and qualitative research methods. Coast *et al.* (2012) argue that qualitative studies are best suited to derive attributes, since they reflect the perspective and experiences of the potential beneficiaries. However, a recent review of the healthcare literature reported that qualitative research methods are being used less often to inform attribute selection, which may make DCEs more susceptible to omitted variable bias if the decision framework is not known prior to the research project (Clark *et al.* 2014).

Creating a choice set and the role of efficiency measures

When creating a choice set, careful consideration needs to be applied in deciding which combinations of attribute levels to present in the questionnaire so attributes can be identified and estimated as precisely as possible. The exact number of questions will depend on the complexity of the particular experiment and the survey approach used. The number of feasible choices will also have an impact on the number of attributes and levels if the full set of combinations is used in a single sample included in the experimental design.

Full factorial and fractional factorial designs

A full factorial design includes all possible combinations of attributes and levels for making profiles or choice sets. Generally, the number of possible profiles is a^n where a is the number of levels and n is the number of possible attributes. If the number of levels varies across attributes then the number of possible hypothetical profiles is $a^n \times b^m$ where a and b are the different attributes levels and n and m are the different attributes.

A full factorial design is likely to yield a very large number of possible choices. For example, in an experiment with two attributes at two levels and four attributes at four levels we have a full factorial of 1204 choices (44*22). Rather than using all the possible combinations, a researcher can opt to use only a fraction of the treatment combinations in what is called a fractional factorial design (FFD) (Louviere *et al.* 2000). FFDs are used to reduce the number of profiles that have to be evaluated, while ensuring enough data are available for statistical analysis, resulting in a carefully controlled set of ‘profiles’ for the respondent to consider. This often results in a large number and experimental design methods are used to create smaller fractional factorial designs. They are extensively documented elsewhere, for example by Reed Johnson *et al.* (2013), and briefly discussed in the paragraphs on orthogonal and D-efficient designs below. If a binary choice DCE is employed (a binary choice is where respondents are presented with a number of profiles (one at a time) and asked how they would choose, with possible responses being 'yes' or 'no'), then the profiles generated from the orthogonal design are the choices. In cases where the respondent is asked to choose between more than one alternative, profiles need to be combined into choice sets.

Creating the choice set

While initial methods for creating choice sets included random pairing of a set of orthogonal choices (e.g. 16 profiles from an orthogonal array would be randomly allocated into eight choices) or the use of a constant comparator (see below), recognition by health economists of the importance of the experimental design component of the DCE has led to the application of two alternative approaches to generating such designs: *orthogonal* and *D-efficient* designs (Reed Johnson *et al.* 2013).

Orthogonal (main effect) design is still the most frequently used approach when creating a DCE to value alternative health care interventions (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012). Software packages (see for example www.spss.com; www.sas.com; www.choice-metrics.com/features.html) are the most common approach to create the orthogonal main effect choice set, and the web-based catalogue approach (e.g. www.york.ac.uk/depts/maths/tables/orthogonal.htm) is still listed as an alternative method to generate a design. Choice sets are then created from singular profiles according to experimental design theory (Reed Johnson *et al.* 2013).

A series of computerised searches have been developed to create choice experiment designs that maximise design efficiency[1] (*D-efficient designs*), and they are easily accessible from computer packages (www.sas.com, www.choice-metrics.com). They are now widely used and present more flexibility than standard tabulated designs from web-based catalogues. Readily available designs are also available from experts (for example forms using methods from Street and Burgess (2007) hosted at the Sydney University of Technology website: http://130.56.248.113/choice/help).

Constant comparator or opt-out option

When a constant comparator is included, each profile within the choice set is compared with the same base comparator (or randomly chosen from the set or predefined according to the present scenario (status quo) or defined as 'neither' of the alternatives presented). The decision to include a base-case scenario or an opt-out alternative should be guided by whether or not the current situation and/or non-participation is a relevant alternative. Not including an opt-out or status quo alternative when this reflects the reality may be problematic as it assumes all respondents choose to use the service and force them in to a choice between two unappealing options, neither of which might have been chosen in practice (Boyle *et al.* 2001).

However, the social care literature shows some cases when respondents can be 'forced' to choose between two separate hypothetical options of care. One example has been reported by Milte and colleagues (2013), where the participants in their study were asked to state their preferences between a pair of hypothetical scenarios reflecting the characteristics of an individualised multidisciplinary rehabilitation programme they would receive at two alternative locations. Given that individuals were already currently participating in a rehabilitation programme a 'forced' choice model was felt appropriate for this application.

In other cases, the choice set includes an additional opt-out option on the respondents' current experience of care. An example is provided by Hall and colleagues (2013) with application to carer support in terminal illness. In the authors' exercise, carers were invited to choose between two hypothetical plans of support services and their current services. Failure to include an opt-out or

---

[1] *Efficiency* indicates the ability of the design to capture information. High efficiency implies that the precision of the parameter estimates is maximised (Hensher *et al*. 2005)

status quo option in such a study risks misrepresenting respondents' choice, because individuals may be 'forced' to choose between plans of support services that do not account for their current services.

The inclusion of the opt-out or status quo option can compromise the statistical design properties of the study. Where the levels for the opt-out or status quo are known in advance the implications for the efficiency of the design can be estimated before collecting data. Where the opt-out or status quo levels are constant across respondents, the effect on design efficiency is smaller. Where information is collected from individuals within the study on their opt out or status quo, and this varies across respondents, the effect on the statistical efficiency of the design is likely to be greater and cannot be estimated in advance of data collection (Ryan 2011).

Designs with restrictions

Sometimes approaches to deriving choice sets may result in implausible combinations (e.g. severe disability with no restrictions on activities of daily life). There are methods for creating designs with restrictions including the nesting approach (Lancsar and Louviere 2008) and use of software packages such as SAS (www.sas.com) and Ngene (www.choice-metrics.com).

Validity of results

Another issue when creating choice sets is whether further choice questions are added to the experimental design, for example to check the validity of the results or as a warm-up exercise. Validity refers to the degree to which a study is able to measure the intended values. Evidence from studies shows that DCEs generally pass these tests of validity (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003). One of the most common strategies for testing validity is the 'dominance test', where one alternative in a choice set is clearly superior to all others (with better outcomes). Responses are said to be 'rational' when the respondent choses this superior alternative.

A few applications in the context of social care used the 'dominance test', for example Ryan *et al.* (2006). Burton *et al.* (2014) also discussed the opportunity of using it, although their survey did not contain any additional choice for 'dominance test' due to the pragmatic limitations of cognitive burden when adding extra choices. DCE guidance reports possible issues in testing 'rational' responses and suggests that it is generally better to include all responses into the final analysis. Deleting respondents may omit valid preferences, leading to biased results (Lancsar and Louviere 2008).

DCE applications may also place emphasis on other validity tests such as: test–retest reliability (using a subsample of respondents to complete a second DCE interview); convergent validity (comparing DCE output with results from other stated preference approaches, e.g. contingent valuation); and theoretical validity (checking that the model coefficients have the expected signs). Details are presented elsewhere (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003).

Splitting the design into separate DCE sets

When a fractional factorial design is not feasible for a participant to complete in one survey, other approaches can be considered. For example, the design can be broken down into blocks. Each block is then given to a different respondent, with the result that the completion of a full design requires several respondents (Hensher *et al.* 2005, Louviere *et al.* 2000). For applications in the social care literature the reader may refer to Burton *et al.* (2014) and Nieboer at al. (2010).

Constructing and conducting the survey

At the start of the questionnaire the DCE choice set needs to be introduced to the respondent by providing information on the project and the experiment itself, explaining the different attributes and levels and presenting an example of a possible answer to a choice. Respondents are presented with a hypothetical situation where they have to imagine the hypothetical services on offer. A general example of such a choice in the social care setting is presented in figure 2.

After completion of the choice task it is possible to include questions to collect information on the respondents' reasons for answering the choice set in the way they did and the degree of difficulty they experienced in answering. These questions can help in identifying possible misunderstandings (Bennett and Blamey 2001). Questions on socio-economic variables, such as age, sex, and income could also be useful to test for heterogeneity in respondents' preferences which could then be accounted for in the modelling.

A piloting of the survey questionnaire is useful in order to test for such issues as readability, the need for rewording, need for additional questions, and overall ease of completion. Information from the pilot phase can inform all the different aspects of DCEs, such as the appropriateness of the attributes and their levels, the length of the choice set to be inserted, the definition of the most suitable design to be used according to design optimality, and the goodness of fit in model estimates. The DCE analysis stage can also be informed by testing the econometric models to be used.

When conducting a DCE different data collection procedures can be employed, including face-to-face interview, telephone interview, mailed questionnaire, email and internet. The face-to-face approach can allow the interviewer to help the respondent to understand the experiment and this has been considered beneficial in some circumstances (Hall *et al.* 2013). However, this has the potential to cause problems of interviewer bias, where the interviewer influences the respondents' responses. Cost and time constraints may limit the use of face-to-face interviews.

Mailed questionnaires have become the preferred option in healthcare research (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012), with applications also in social care (for examples see Mentzakis *et al.* (2011) and Nieboer *et al.* (2010)). A telephone interview is a cheaper option than a face-to-face one, but there are still issues related to the need for the respondent to read the questionnaire in order to better understand the task. Paper or electronic surveys are, thus, preferred. The paper mailed approach is commonly used in health economics, being less expensive than others and guaranteeing the respondent dedicated time to complete it (Mentzakis *et al.* 2011). Electronic approaches are gaining recognition due to the fact that they allow for the collection of large data sets relatively cheaply and quickly (Clark *et al.* 2014, Nieboer *et al.* 2010). They also allow for the presentation of more complex choice/questionnaire designs not otherwise easily employable, such as where subsequent choices are defined according to what has been previously chosen (Cairns and van der Pol 2004). Additional developments (for both paper and online surveys) include the opportunity to include pictorial representations of the domains to facilitate comprehension of the questions. One example from the social care literature is given by Negrín *et al.* (2008). It showed that the application of pictorial representation of the attributes' levels is able to facilitate engagement with the respondents and minimise cognitive burden in the choice process.

Sampling in discrete choice modelling depends on the following considerations:

- level of accuracy (precision) required – in general, the higher the level of accuracy needed, the larger the sample size needs to be;
- planned subgroups analysis in addition to the overall population level analysis – the overall sample size needs to be large enough that an adequate level of accuracy of analysis for these subgroups can also be achieved;

- the data collection method to be used – self-completion questionnaires are often cheaper compared with other collection methods such as interviews, but their response rates tend to be lower, so the sample size should be larger; and
- the level of variability between responses – the less variable the responses are the smaller is the sample size required to achieve the same level of accuracy (Ryan et al. 2008).

Although sample size calculation is an ill-defined area within DCEs, different approaches are available from the literature, and they match the following sampling strategies:

- simple random samples, based on the desired level of accuracy of the estimated probabilities;
- stratified random samples, where the population is first divided into mutually exclusive groups relevant to the topic of the study and then a random sample is drawn from each stratum; and
- choice-based sample, where the sample is selected to allow all alternatives to be chosen by a significant proportion of sampled decision-makers (Hensher et al. 2005; Louviere et al. 2000).

The strengths and weaknesses of these approaches are extensively discussed elsewhere (Daniel 2012). Social care applications may also use past studies as guidance for their experiments or refer to the average sample size reported in previous literature (de Bekker-Grob *et al.* 2012).

Analysis of data and interpretation of results

Analysis of data

Data collected from a discrete choice experiment should be analysed by regression analysis where the dependent variable takes on only two possible values (chosen, not chosen). The analysis of data involves estimating a utility or satisfaction function which specifies the relationship between the service attributes and preferences using regression analysis.

When the choice faced by respondents in a DCE is binary or if the choice set includes only two ‘forced’ alternatives (i.e. would you choose alternative A or B), binary logit and probit are the commonly used models in social care. Few studies in health and social care focus on binary choice or ‘forced’ choice models. For examples the reader can refer to Burton *et al.* (2014), examining preferences for support interventions after stroke, or Dixon *et al.* (2013), investigating preferences for the delivery of different community-based models of care.

As more studies recognised the importance of allowing for multiple options, multinomial rather than binary choice data were collected (e.g. by including neither or opt-out options). The Conditional Logit Model (CLM) has been employed as a standard approach to estimating multinomial choice models with the assumption that options available are perfect substitutes. Although respondents to a DCE are presented with a set of multiple observations, no serial correlation within a respondent’s series of responses is assumed. This assumption might be violated where, for example, respondents learn how to answer the DCE questionnaire from their previous responses, or they suffer from questionnaire fatigue. Another issue is that CLM implies constant observed factors for each individual, regardless of any known or unknown characteristics. This type of preference variation can be incorporated into the modelling process by interacting individual known characteristics with the attributes included in the DCE. Several CLM applications are reported in the social care literature (e.g. Burge *et al.* 2010, Hall *et al.* 2013, Hoefman *et al.* 2014) .

More advanced models for analysis are also available to relax the restrictions of the CLM (e.g. the nested logit, the mixed logit or random parameter logit, and latent class), and their characteristics are extensively commented on elsewhere (Ryan *et al.* 2008). Their application to social care is relatively new (Burge *et al.* 2010, Mentzakis *et al.* 2011, Negrín *et al.* 2008, Netten *et al.* 2012).

Interpretation of results

Raw preferences can be presented to see how choices are split between different alternatives. The total number of raw preferences is calculated from the number of choices per set multiplied by the number of respondents. Percentages of preferences across alternatives can be calculated from the frequencies of choice for each alternative (Tinelli *et al.* 2009).

Consideration can be given here to whether or not any individual always chooses a given option. First, respondents could always choose option A over B (or vice versa), or the first option rather than the second in each choice set. Secondly, they could also present constant preferences for a particular option on offer, e.g. opt-out (current situation/no service) versus alternative services on offer. In such circumstances, it would be unclear if they were still trading across alternatives when facing each choice of the DCE experiment. Subgroup analysis across respondents could be employed to compare preferences across groups of subjects with or without constant preferences (Tinelli *et al.* 2009).

A third case of constant preferences occurs when only one attribute matters to the choice (lexicographic preferences). Some respondents (respondents with 'dominant preferences') consistently chose the scenario with the 'best' level of a particular dimension. This may arise both from the complexity of the DCE choices but also an individual's past experiences or expectations (Scott 2002).

From the coefficient estimates it is possible to identify whether the attribute influences the preferences for that particular good or service, the relative importance of individual attributes, the trade-offs or marginal rates of substitution between these attributes and the overall utility from a scenario.

The sign, relative size and significance level of the regression coefficients show the relative importance of the different attributes to individual preferences. A positive sign on the coefficient indicates that the higher that attribute level, the higher the level of utility derived. The relative size of the coefficients indicates the relative importance of that attribute in determining the overall utility, taking into account the different units of measurement considered. However, it gives no indication of the strength of preferences, which is why the marginal rate of substitution (MRS) is calculated, this being the rate at which respondents are willing to give up one attribute for another (where all the rest are constant). It is worth noting that estimation of trade-offs can only be conducted when a continuous variable is included. This continuous variable is most commonly price (e.g. monetary compensation to conduct carer support tasks (see Mentzakis *et al.* (2011)), but risk and time have also been used in social care applications of DCE (e.g. risk of falling and the duration of effort required in the rehabilitation session) (Milte *et al.* 2013).

Another objective might be estimating how preferences vary by individual respondent characteristics. For policy analysis, researchers may calculate how choice probabilities vary with changes in attributes or attribute levels or calculate secondary estimates of money equivalence (willingness to be compensated for caregiver task (Mentzakis *et al.* 2011)), risk equivalence (maximum acceptable risk of falling) or time equivalence (maximum acceptable duration of effort required in the rehabilitation session) for changes in scenarios (Milte *et al.* 2013).

DCEs and their application to social care

Discrete choice experiments have been widely applied to health care settings in different contexts. These include: eliciting patient/community preferences in the delivery of health services; establishing consultants' preferences when setting priorities across interventions; developing outcome measures; eliciting patient preferences in the doctor-patient relationship; and evaluating alternatives within a randomised controlled trial (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003). Given that DCEs are an established health services research approach, a comprehensive search of the literature has been conducted to identify examples of their application to social care.

Literature search

The focus of the search was to identify published English language studies using DCEs within the social care context. The definition of social care applied for the literature search was presented in the introduction.

After consultation with a librarian at the LSE searches of the following electronic databases from 1 January 2000 to 31 December 2014 were undertaken: Social Care Online, EBSCO, ISI Social Science, Pubmed, Proquest and Scopus.

Papers were screened by the author for DCEs adopting a search strategy presented elsewhere (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003). The search terms included: 'discrete choice experiment(s)', 'discrete choice model(l)ing', 'stated preference', 'part-worth utilities', 'functional measurement', 'paired comparisons', 'pairwise choices', 'conjoint analysis', 'conjoint measurement', 'conjoint studies', or 'conjoint choice experiment(s)'. Applications to social care setting were extrapolated using the terms: 'social care', 'social services', 'carer social support', 'caregiver(s)', or 'informal care'.

Background results

A total of 372 potentially relevant papers were identified. Following title and abstract screening a total of 12 papers were selected after excluding evidence not meeting the inclusion criteria or duplicates (Burge *et al.* 2010, Burton *et al.* 2014, Dixon *et al.* 2013, Hall *et al.* 2013, Hoefman *et al.* 2014, Mentzakis *et al.* 2011, Milte *et al.* 2013, Negrín *et al.* 2008, Netten *et al.* 2012, Nieboer *et al.* 2010, Potoglou *et al.* 2011, Ryan *et al.* 2006). Alternative formats of choice modelling from the discrete choice experiment (such as rating conjoint analysis (van den Berg *et al.* 2008, 2005) or best-worst scaling (Louviere *et al.* 2010)) were not considered in this analysis of the literature.

Overall the picture is of the application of DCE to social care being quite recent (table 1). The first application appeared in the literature in 2006 (Ryan *et al.* 2006), and about half of the subsequent papers have been published since 2012. Ten of publications reported on research based in Europe: seven from the UK, followed by the Netherlands (two) and Spain (one)); another two came from Australia.

Many of the applications (5/12) were used to develop and estimate weights for social care outcome measures. They included: older persons' utility scale (OPUS (Ryan *et al.*, 2006)), ASCOT (Burge *et al.* 2010, Netten *et al.* 2012, Potoglou *et al.* 2011) and CareQol (Hoefman *et al.* 2014). Four papers were aimed at valuing social care interventions for older people and three at valuing carer support.

Results from the review are reported and discussed in Table 1, looking at the four key stages of a DCE (i.e. attributes and their levels, creating a choice set, constructing and conducting the survey, and analysis of data and interpretation of findings).

TABLE 1 : Literature search – description of the DCE studies

	Number of papers	References
Date of publication		
Before 2012	6	Burge et al. (2010), Mentzakis et al. (2011), Negrín et al. (2008), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
2012–2014	6	Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Milte et al. (2013), Netten et al. (2012)
Country of origin		
UK	7	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
The Netherlands	2	Hoefman et al. (2014)
Spain	1	Negrín et al. (2008)
Australia	2	Hall et al. (2013), Milte et al. (2013)
Purpose of the paper		
Developing and estimating weights for social care outcome measures	5	Burge et al. (2010), Hall et al. (2013), Milte et al. (2013), Netten et al. (2012), Potoglou et al. (2011), Ryan et al. (2006)
Valuing social care interventions for older people	4	Dixon et al. (2013), Milte et al. (2013), Negrín et al. (2008), Nieboer et al. (2010)
Valuing carer support	3	Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011)

Establishing attributes and their levels

In previous social care DCEs, attributes and their levels were mainly derived from previous research, policy documents or previous experiments (75% of cases). Additional papers (42%) reported using qualitative approches, as suggested by Coast *et al.* (2012). More details are reported in Table 2.

The number of attributes varied between four and 14 (with a mean of seven). According to a recent review of the healthcare evidence (Clark *et al.* 2014), a past trend towards including fewer attributes seems to be reversing, with about half of the applications reporting between eight and 14 attributes. In studies with a large number of attributes the original set of eight to 14 attributes was split across two separate DCE surveys to be run in parallel within the same study (5/12). This allowed building manageable experiments in line with current guidance.

Overall, the attribute levels varied between two and five, with the majority of applications reporting on attributes with two or four levels only (8/12). Similar results are available from healthcare studies (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012).

TABLE 2 : Results of the literature search – attributes and their levels

	Number of papers	References
Sources		
Literature, policy documents or previous experiments	9	Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
Qualitative interviews/focus groups or workshops	5	Burge et al. (2010), Burton et al. (2014), Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011)
Views of experts	4	Milte et al. (2013), Netten et al. (2012), Potoglou et al. (2011), Ryan et al. (2006)
Piloting with study participants	2	Burge et al. (2010), Mentzakis et al. (2011)
Number of attributes	4 and 14 (min, max)	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
Number of levels	2 and 5 (min, max)	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
Attributes were split across two separate DCE sets	5	Burge et al. (2010), Burton et al. (2014), Hall et al. (2013), Netten et al. (2012), Potoglou et al. (2011)

Creating a choice set and the role of efficiency measures

When building the choice set design in the social care studies, issues were carefully considered to reflect recent trends in healthcare research literature concerning the use of more flexible and appropriate DCE designs and the inclusion of opt-out options (Clark *et al.* 2014). Half the applications adopted orthogonal designs available from catalogues or derived from software, while the other half used optimal design procedures (Table 3). Three papers reported on choices between two hypothetical alternatives (3/12). The majority of the applications (8/12) addressed issues related to opt-out options to include one additional alternative: being 'not sure', for example see Burge *et al.* (2010); 'no care provided'; for example see Mentzakis *et al.* (2011); 'current care', for example see Hall *et al.* (2013) and Negrín *et al.* (2008).

A few applications of DCE in social care research added additional choice options to test for validity of results using a dominance test (see Burge *et al.* (2010) and Ryan *et al.* (2006)) or discussed the opportunity of testing for validity (Burton *et al.* 2014). More details on validity tests are discussed in the results section below.

The number of choice sets in the experimental design varied between 14 and 256. In order to reduce the number of choices in the set to a more manageable number, almost all of the studies (11/12) split the DCE into separate blocks, with a range between five and 16 choices. A similar pattern has been reported in healthcare research (Clark *et al.* 2014, de Bekker-Grob *et al.* 2012).

TABLE 3 : Results of the literature search – creating the choice set

	Number of papers	%	References
Design type			
Fractional factorial	12	100	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)
Design method			
Orthogonal designs (catalogues or derived from software)	6	50	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hoefman et al. (2014), Nieboer et al. (2010), Ryan et al. (2006)
Optimal designs ('Huber and Zwerina designs' or 'Street and Burgess designs')	6	50	Hall et al. (2013), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011)
Number of choices			
2	3	25	Burton et al. (2014), Dixon et al. (2013), Ryan et al. (2006)
3	9	75	Burge et al. (2010), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011)
Splitting the DCE into separate DCE sets			
Blocks (of up to 16 choices)	11	92	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)

Constructing and conducting the survey

In the majority of the cases the development of the questionnaire was informed by piloting with a subset of respondents before starting data collection (7/12; see Table 4). Similar results were reported elsewhere in the healthcare research literature (Mandeville *et al.* 2014). Unfortunately, there is no clear guidance available on sample size nor sampling strategy for DCEs, and half of the social care studies did not report any specification on their calculation for these. The remaining studies referred to: simple random sampling approach (2); stratified random sampling approach (2) ; and previous cases from the literature (2). The target samples included: general adult population (5/12), older people (3/12), patients/service users (2/12), or carer (3/12). Burton *et al.* (2014) compared preferences results between different groups of patients and their carers.

Face-to-face interview was the preferred data collection method (8/12) followed by mailed questionnaire (4/12). Again, the proportions are similar in healthcare research (de Bekker-Grob *et al.* 2012). Two articles included pictorial representations of the attributes rather than written

TABLE 4 : Results of the literature search – constructing and conducting the survey

	Number of papers	References
Piloting	7	Burge et al. (2010), Hall et al. (2013), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011), Ryan et al. (2006)
Sample size calculations		
Simple random samples	2	Burge et al. (2010), Negrín et al. (2008)
Stratified random samples	2	Nieboer et al. (2010), Potoglou et al. (2011)
Previous examples from the literature	2	Burton et al. (2014), Hall et al. (2013)
Not specified/reported elsewhere	6	Dixon et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Netten et al. (2012), Ryan et al. (2006)
Target sample		
General adult population	5	Burge et al. (2010), Hoefman et al. (2014), Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011)
Older people	3	Milte et al. (2013), Nieboer et al. (2010), Ryan et al. (2006)
Patients	2	Burton et al. (2014), Dixon et al. (2013)
Carers	3	Burton et al. (2014), Hall et al. (2013), Mentzakis et al. (2011)
Data collection method		
Face-to-face interviews	8	Burge et al. (2010), Dixon et al. (2013), Hall et al. (2013), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011), Ryan et al. (2006)
Paper questionnaires	2	Burton et al. (2014), Mentzakis et al. (2011)
Online questionnaires	2	Hoefman et al. (2014), Nieboer et al. (2010)
Description of the choices		
Pictorial representation	2	Burton et al. (2014), Negrín et al. (2008)
In writing	10	Burge et al. (2010), Burton et al. (2014), Dixon et al. (2013), Hall et al. (2013), Hoefman et al. (2014), Mentzakis et al. (2011), Milte et al. (2013), Negrín et al. (2008), Netten et al. (2012), Nieboer et al. (2010), Potoglou et al. (2011), Ryan et al. (2006)

descriptions, which were used in the remaining ten papers. The review by de Bekker-Grob *et al.* (2012) reported little progress on this matter in healthcare research, with only two studies in their review using pictures/figures to help respondents' understanding of risk.

Analysis of data and interpretation of findings

In the literature included here, responses from choices were modelled using random utility framework (Ryan *et al.* 2008). Random effect logit/probit models were employed for binary choice models (3/12). Multinomial logit/conditional logit were the preferred models of analysis for multiple choice experiments (5/12). More advanced models for analysis were also used to relax multinomial logit restrictions, including nested logit (1/12), mixed logit (3/12) or latent class (1/12) models (see Table 5). One article used both multinomial logit and nested logit models to inform the analysis framework (Burge *et al.* 2010). The trend towards using more flexible econometric models in DCE evaluations, including mixed logit and latent class, remains stronger in the healthcare research literature (Clark *et al.* 2014).

TABLE 5 : Results of the literature search – analysis of data and interpretation of findings

	Number of papers	References
Models of analysis		
multinomial logit/conditional logit	5	Burge et al. (2010), Hall et al. (2013), Hoefman et al. (2014), Milte et al. (2013), Nieboer et al. (2010)
nested logit	1	Burge et al. (2010)
random effect logit/probit models	3	Burton et al. (2014), Dixon et al. (2013), Ryan et al. (2006)
mixed logit	3	Negrín et al. (2008), Netten et al. (2012), Potoglou et al. (2011)
latent class	1	Mentzakis et al. (2011)
MRS		
cost	4	Burge et al. (2010), Mentzakis et al. (2011); Negrín et al. (2008), Nieboer et al. (2010)
time	2	Burton et al. (2014), Milte et al. (2013)
risk	1	Milte et al. (2013)
Validity of results		
rationality (dominance) test	1	Ryan et al. (2006)
reliability test	2	Burge et al. (2010), Ryan et al. (2006)
convergent validity test	1	Ryan et al. (2006)
theoretical test	1	Mentzakis et al. (2011)

All papers in this review reported on the parameter estimates from the regression analysis, and commented on which attributes influenced the respondent preferences for the service considered and their relative importance. In line with current trends in healthcare research, where reporting of monetary values has fallen compared with earlier periods of research (Clark *et al.* 2014), only a few papers (4/12) included a cost attribute and calculated respondent willingness to be compensated for unit increases in a particular attribute. For example, Mentzakis *et al.* (2011) reported that caregivers were willing to accept a payment of £0.38 to £0.83 per hour for the personal care service they provided. Additional studies calculated MRS between attributes and time or attribute and risk. For example, Milte *et al.* (2013) reported that respondents were prepared to accept an increase in the duration of the rehabilitation programme of over two hours in a two-month period (132 minutes) in order to achieve the highest mobility outcome of walking with a stick unaided; or a 22% increase in the risk of falling and breaking another bone to avoid enduring severe pain from participating in a rehabilitation programme.

In contrast with current practice in healthcare research (Clark *et al.* 2014), only a few articles tested the validity of results (3/12). The tests adopted by Ryan *et al.* (2006) included dominance tests (inconsistent responses were dropped), test–retest reliability (a subsample completed a second DCE interview) and convergent validity of quality weights (respondents' rating of 16 vignettes). In another application Burge *et al.* (2010) deleted responses if they did not pass accuracy checks. One additional paper considered the theoretical validity of the results (Mentzakis *et al.* 2011). When testing for irrational/inconsistent responses none of the applications included comparison between all responses and subsets without irrational/inconsistent responses in the final analysis, as suggested by current guidance (Lancsar and Louviere 2008), but instead they deleted all irrational/inconsistent responses from their analysis (Burge *et al.* 2010, Ryan *et al.* 2006).

Recommendations

for further development of DCEs to support their use in social care research

DCEs have a strong history of use in health economics and in recent years they have supported the evaluation of social care interventions as well as the development of social care-related quality of life and wellbeing measures. However, the application of DCEs to social care is far from straightforward and a number of challenges have been raised in the literature.

The cognitive burden to respondents

There is discussion around the fact that DCE designs could put too much cognitive burden on some respondents, and participants could react by opting for decisions that do not reflect their true preferences (Bilger *et al.* 2013, Burton *et al.* 2014, Negrín *et al.* 2008, Netten *et al.* 2012, Whitty *et al.* 2013). This is particularly challenging when dealing with social care evaluations, where the target sample of the research is more likely to include individuals in need or at risk, or individuals with needs arising from illness, disability, old age or poverty. A few related recommendations for future research are presented below.

Selecting the attributes and levels

The construction of the choice sets included in an experiment requires a careful selection of the attributes and levels to be used. When respondents are asked to trade-off between a large number of attributes, one option could be to concentrate on a subset of attributes for the DCE. However, this would limit the findings of the study and result in biased estimates. A solution reported in the social care literature is to split the attributes across separate DCE sets with overlap in some attributes to allow presentation and results to be based on a common utility scale. This would allow the full set of domains to be considered (Burge *et al.* 2010, Hall *et al.* 2013, Potoglou *et al.* 2011). Rather than conducting this as two separate studies each looking at one of the subset of domains, each respondent in the DCE survey can be asked to complete both experiments, to allow the samples for the two sub-studies to be matched (Netten *et al.* 2012).

Creating the choice set

The second main step in the experimental design involves the construction of the choice combinations. There is evidence that task complexity may increase with the number of choice sets beyond a certain threshold (Bech *et al.* 2011). Decisions on size of choice set may require trading parameter estimability, efficiency and precision of estimates against the cognitive burden of the experiment on the respondents (Lancsar and Louviere 2008). Experimental design techniques can be helpful in defining the appropriate choice set for application to social care. A DCE application evaluating end of life care demonstrated that recent advances in experimental design techniques made it possible to compare different design approaches and showed that the use of less efficient designs may be justified when choice tasks are complex (Bilger *et al.* 2013).

Splitting the DCE into several DCE sets

A successful approach applied in the social care literature is to divide the DCE into different segments or blocks. The blocking approach helped by reducing the number of choices and enabled researchers to keep the questionnaire to a more manageable size (Bridges *et al.* 2011, Reed Johnson *et al.* 2013).

Using pictorial representations of the subject

Different formats are available to the researchers when presenting the choice sets to the respondents, including pictorial representations versus written format (de Bekker-Grob *et al.* 2012, Ryan and Gerard 2003). Only a few DCE applications used pictures and figures to represent the attributes and their levels in a social care setting, but they proved to be successful and facilitated comprehension of the questions by respondents (Burton *et al.* 2014, Negrín *et al.* 2008). Future work should further explore the benefit of using pictorial representations versus written format in social care.

Collecting the data via interviews

Although requiring more time and resources, face-to-face interviews with a trained researcher can provide direct support to the respondents and help them to better understand the task. This particular approach appeared to be the preferred data collection method reported in the social care literature (Burge *et al.* 2010, Dixon *et al.* 2013, Hall *et al.* 2013, Milte *et al.* 2013, Negrín *et al.* 2008, Netten *et al.* 2012, Potoglou *et al.* 2011, Ryan *et al.* 2006). Exploring whether different methods of data collection work better in different settings and/or with different client groups would also be valuable.

Choice of modelling for analysis

An additional key area for discussion in the DCE literature concerns the analysis of results. From the literature review reported here we are aware that the majority of the DCE applications to social care included an opt-out option within the choice set and adopted a DCE framework with multiple choice alternatives. Conditional logit/multinomial logit appeared to be the most common method for analysing DCE data (Burge *et al.* 2010, Hall *et al.* 2013, Hoefman *et al.* 2014, Milte *et al.* 2013, Nieboer *et al.* 2010). However, the review identified also a number of more sophisticated analytical approaches, including nested logit (Burge *et al.* 2010), mixed-logit (Negrín *et al.* 2008, Netten *et al.* 2012, Potoglou *et al.* 2011) and latent class (Mentzakis *et al.* 2011). The latter approaches are able to drop some of the assumptions of the conditional logit/multinomial logit and allow for investigating the extent to which (unobserved) individual preferences for a service vary across respondents. However, their application to policy is limited as it is not possible to identify where preferences differ (Ryan 2011). It could be argued that a more appropriate approach could be to adopt conditional logit and apply it as (alternative) model(s) to gain better insight into observed variation that can potentially be used by policymakers.

External validity of responses

It is argued that the greatest challenge facing practitioners of DCE is to test the external validity of responses; that is, the extent to which respondents behave in reality as they state in a hypothetical context (Lancsar and Swait 2014, Ryan *et al.* 2008). Unfortunately, there is little research available on this challenge in healthcare (de Bekker-Grob *et al.* 2012, Lancsar and Swait 2014, Ryan and Gerard 2003), probably reflecting the difficulty of investigating this question in the UK publicly-provided health care system where individuals have limited choice and do not pay at the point of consumption. The context for social care is often quite different, which may make it more amenable to real-world research on choices.

Crucially such research would help provide further guidance on good practice for DCEs, and more credibility and confidence in results from well-conducted DCE studies that follow good research practice (Bridges *et al.* 2011, Reed Johnson *et al.* 2013). The issue remains in social care research applications of DCE, where it is still unclear whether people would pay as much in real life as they claim they would (Nieboer *et al.* 2010).

Using alternative stated preference approaches. When creating the choice set a focus of attention in the literature is the inclusion of a base-case scenario (status quo or no service) as an added comparator within the choice set if the current situation and/or non-participation is a relevant alternative (Bridges *et al.* 2011, Reed Johnson *et al.* 2013). However, keeping two or more alternatives in mind at once may be a difficult task, particularly for people with some cognitive problems (if not assisted in the choice process, see above). An alternative stated preferences approach, best–worst scaling (BWS), has been already introduced to social care research (Flynn *et al.* 2007, Koopmanschap *et al.* 2008, Marley *et al.* 2008, Potoglou *et al.* 2011, van den Berg *et al.* 2004). Instead of respondents making choices between alternatives, they are presented with alternatives one at a time and make choices within each of them. BWS appears to facilitate collection of preference data from service users as it is likely to secure lower cognitive burden compared with DCE, where keeping two or more profiles in mind is likely to be a harder task.

Using BWS in social care research has had success but it remains a relatively new method, and more comparative research is needed to demonstrate the superiority of BWS over DCE in terms of respondents' fatigue and responses, and the robustness of the resulting modelling estimates across settings (Whitty *et al.* 2013).

There is evidence to support the application of DCE as a useful tool to inform the evaluation of social care interventions as well as the development of social care-related quality of life and wellbeing measures. Past DCE applications can provide helpful guidance to social care researchers, policymakers and practitioners when valuing preferences in social care settings, and they are summarised in this review. However caution should be exercised when applying DCE to social care, as important issues need to be addressed, including the cognitive burden to respondents, choice of modelling for analysis, and external validity of responses. Alternative stated preferences approaches (see above for details on BWS) are also available and their application to social care should be further explored.

References

Bech M, Kjaer T, Lauridsen J (2011) Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment, *Health Economics*, 20, 3, 273–286.

Bennett J, Blamey R (eds) (2001) *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar, Cheltenham.

Bilger M, Flynn T, Farooqui A, Malhotra C, Finkelstein E (2013) Cognitive burden in discrete choice experiments: the case of preferences for end-of-life in Singapore, presentation at International Choice Modelling Conference, Sydney, icmconference.org.uk/index.php/icmc/ICMC2013/paper/viewFile/543/333.

Boyle KJ, Holmes TP, Teisl MF, Roe B (2001) A comparison of conjoint analysis response formats, *American Journal of Agricultural Economics*, 83, 2, 441–454.

Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, Johnson FR, Mauskopf J (2011) Conjoint analysis applications in health – a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force, *Value in Health*, 14, 4, 403–413.

Burge P, Netten A, Gallo F (2010) Estimating the value of social care, *Journal of Health Economics*, 29, 6, 883–894.

Burton CR, Fargher E, Plumpton C, Roberts GW, Owen H, Roberts E (2014) Investigating preferences for support with life after stroke: a discrete choice experiment, *BMC Health Services Research*, 14, 63.

Cairns J, van der Pol M (2004) Repeated follow-up as a method for reducing non-trading behaviour in discrete choice experiments, *Social Science & Medicine*, 58, 11, 2211–2218.

Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW (2014) Discrete choice experiments in health economics: a review of the literature, *PharmacoEconomics*, 32, 9, 883–902.

Coast J, Al-Janabi H, Sutton EJ, Horrocks SA, Vosper AJ, Swancutt DR, Flynn TN (2012) Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations, *Health Economics*, 21, 6, 730–741.

Daniel J (2012) *Sampling Essentials. Practical Guidelines for Making Sampling Choices*, Sage, London.

de Bekker-Grob EW, Ryan M, Gerard K (2012) Discrete choice experiments in health economics: a review of the literature, *Health Economics*, 21, 2, 145–172.

Dixon S, Nancarrow SA, Enderby PM, Moran AM, Parker SG (2013) Assessing patient preferences for the delivery of different community-based models of care using a discrete choice experiment, *Health Expectations*, 1–11.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005) *Methods for the Economic Evaluation of Health Care Programme. Third Edition – Research Database*, University of York, York.

Faria R, Weatherly H, Van Den Berg B (2012) A review of approaches to measure and monetarily value informal care, in Curtis, L (ed.), *Unit Costs of Health and Social Care 2012*, PSSRU, Canterbury, pp. 22–31.

Flynn TN, Louviere JJ, Peters TJ, Coast J (2007) Best-worst scaling: what it can do for health care research and how to do it, *Journal of Health Economics*, 26, 1, 171–189.

Grutters JPC, Kessels AGH, Dirksen CD, van Helvoort-Postulart D, Anteunis LJC, Joore MA (2008) Willingness to accept versus willingness to pay in a discrete choice experiment, *Value in Health*, 11, 7, 1110–1119.

Hall J, Kenny P, Hossain I, Street DJ, Knox SA (2013) Providing informal care in terminal illness: an analysis of preferences for support using a discrete choice experiment, *Medical Decision Making*, 34, 6, 731–745.

Hensher DA, Rose JM, Greene WH (2005) *Applied Choice Analysis Primer*, Cambridge University Press, Cambridge.

Hoefman RJ, van Exel J, Rose JM, van de Wetering EJ, Brouwer WBF (2014) A discrete choice experiment to obtain a tariff for valuing informal care situations measured with the CarerQol instrument, *Medical Decision Making*, 34, 1, 84–96.

Koopmanschap MA, van Exel JNA, van den Berg B, Brouwer WBF (2008) An overview of methods and applications to value informal care in economic evaluations of healthcare, *PharmacoEconomics*, 26, 4, 269–280.

Lancsar E, Louviere J (2008) Conducting discrete choice experiments to inform healthcare decision making: a user's guide, *PharmacoEconomics*, 26, 8, 661–677.

Lancsar E, Swait J (2014) Reconceptualising the external validity of discrete choice experiments, *PharmacoEconomics*, 32, 10, 951–965.

Louviere JJ, Flynn TN, Carson RT (2010) Discrete choice experiments are not conjoint analysis, *Journal of Choice Modelling*, 3, 3, 57–72.

Louviere JJ, Hensher DA, Swait JD (2000) *Stated Choice Methods: Analysis and Applications*, Cambridge University Press, Cambridge.

Mandeville KL, Lagarde M, Hanson K (2014) The use of discrete choice experiments to inform health workforce policy: a systematic review, *BMC Health Services Research*, 14, 1, 367.

Marley AAJ, Flynn TN, Louviere JJ (2008) Probabilistic models of set-dependent and attribute-level best–worst choice, *Journal of Mathematical Psychology*, 52, 5, 281–296.

Marshall D, Bridges JFP, Hauber B, Cameron R, Donnalley L, Fyie K, Reed Johnson F (2010) Conjoint analysis applications in health — how are studies being designed and reported? *The Patient: Patient-Centered Outcomes Research*, 3, 4, 249–256.

Mentzakis E, Ryan M, McNamee P (2011) Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity, *Health Economics*, 20, 8, 930–944..

Milte R, Ratcliffe J, Miller M, Whitehead C, Cameron ID, Crotty M (2013) What are frail older people prepared to endure to achieve improved mobility following hip fracture? A discrete choice experiment, *Journal of Rehabilitation Medicine*, 45, 1, 81–86.

National Institute for Health and Care Excellence (2013) *The Social Care Guidance Manual*, NICE, Manchester, www.nice.org.uk/article/pmg10/resources/non-guidance-the-social-care-guidance-manual-pdf.

Negrín MA, Pinilla J, León CJ (2008) Willingness to pay for alternative policies for patients with Alzheimer's Disease, *Health Economics, Policy, and Law*, 3, 3, 257–275.

Netten A (2011) *Overview of Outcome Measurement for Adults Using Social Care Services and Support*, NIHR SSCR Methods Review 6, NIHR School for Social Care Research, London.

Netten A, Burge P, Malley J, Potoglou D, Towers A-M, Brazier J, Flynn T, Forder J, Wall B (2012) Outcomes of social care for adults: developing a preference-weighted measure, *Health Technology Assessment*, 16, 16, 1–166.

Nieboer AP, Koolman X, Stolk EA (2010) Preferences for long-term care services: willingness to pay estimates derived from a discrete choice experiment, *Social Science & Medicine (1982)*, 70, 9, 1317–1325.

Oremus M, Tarride J-E (2008) A systematic review of the use of contingent valuation in Alzheimer's disease research, *Dementia*, 7, 4, 461–480.

Potoglou D, Burge P, Flynn T, Netten A, Malley J, Forder J, Brazier JE (2011) Best-worst scaling vs. discrete choice experiments: an empirical comparison using social care data, *Social Science & Medicine*, 72, 10, 1717–1727.

Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, Bresnahan BW, Kanninen B, Bridges JFP (2013) Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force, *Value in Health*, 16, 1, 3–13.

Ryan M (2011) Taking conjoint analysis to task, *Value in Health*, 14, 4, 401–402.

Ryan M, Gerard K (2003) Using discrete choice experiments to value health care programmes: current practice and future research reflections, *Applied Health Economics and Health Policy*, 2, 1, 55–64.

Ryan M, Gerard K, Amaya-Amaya M (2008) *Using Discrete Choice Experiments to Value Health and Health Care*, Springer, Dordrecht.

Ryan M, Netten A, Skåtun D, Smith P (2006) Using discrete choice experiments to estimate a preference-based measure of outcome – an application to social care for older people, *Journal of Health Economics*, 25, 5, 927–944.

SCIE, Francis J, Byford S (2011) *SCIE's Approach to Economic Evaluation in Social Care*, Social Care Institute for Excellence, London.

Scott A (2002) Identifying and analysing dominant preferences in discrete choice experiments: an application in health care, *Journal of Economic Psychology*, 23, 383–398.

Street DJ, Burgess L (2007) *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, John Wiley, Chichester.

Tinelli M (2007) *Developing and applying discrete choice experiments (DCEs) to inform pharmacy policy*, PhD Thesis, University of Aberdeen

Tinelli M, Ryan M, Bond C (2009) Patients' preferences for an increased pharmacist role in the management of drug therapy, *International Journal of Pharmacy Practice*, 17, 5, 275–282.

van den Berg B, Al M, van Exel J, Koopmanschap M, Brouwer W (2008) Economic valuation of informal care: conjoint analysis applied in a heterogeneous population of informal caregivers, *Value in Health*, 11, 7, 1041–1050.

van den Berg B, Bleichrodt H, Eeckhoudt L (2005) The economic value of informal care: a study of informal caregivers' and patients' willingness to pay and willingness to accept for informal care, *Health Economics*, 14, 4, 363–376.

van den Berg B, Brouwer WBF, Koopmanschap MA (2004) Economic valuation of informal care. An overview of methods and applications, *European Journal of Health Economics*, 5, 1, 36–45.

Whitty J, Ratcliffe J, Chen G, Scuffham P (2013) A comparison of discrete choice and best worst scaling methods to assess Australian public preferences for the funding of new health technologies, presentation at International Choice Modelling Conference, Sydney, www.icmconference.org.uk/index.php/icmc/ICMC2013/paper/viewFile/612/286.