

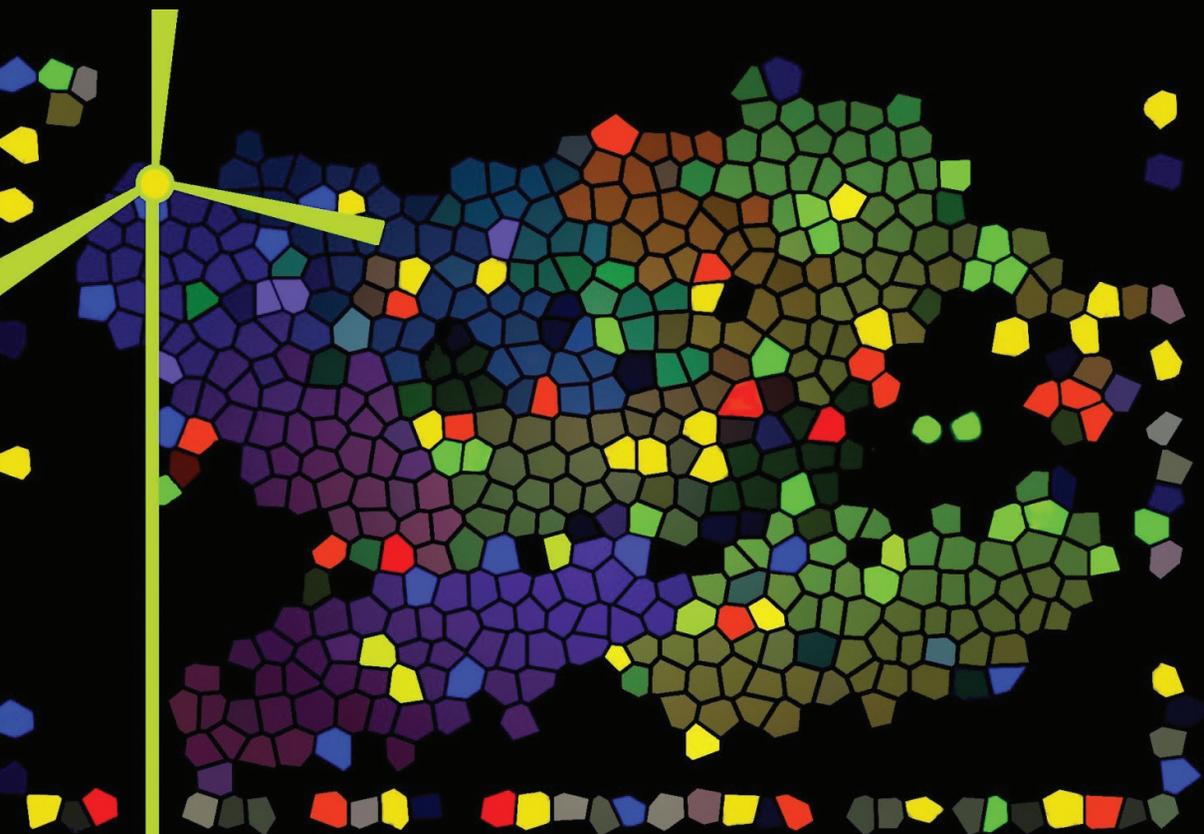
IMPRESA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

ASSESSMENT METHODOLOGIES

ENERGY, MOBILITY AND OTHER
REAL WORLD APPLICATION

PEDRO GODINHO
JOANA DIAS

EDITORS



PREFERENCE BASED DECISIONS IN HEALTH

Patrícia Antunes^{1,4}, Pedro Lopes Ferreira^{2,4} and Lara Noronha Ferreira^{3,4}

Abstract

The decision making process within a context of limited resources has important consequences on the health of individuals. The criteria by which health technology assessment and health interventions are decided, as well as the review of current procedures represent a challenge for decision makers in the face of the current economic slowdown. The decision making process requires a transparent and rational approach to ensure the maximization of health care and the reduction of inequalities. Economic evaluation techniques may provide a valid solution to achieve these goals. Preference based measures are part of economic evaluation techniques. These approaches incorporate utilities for health outcomes and can be used in cost-benefit analyses to aid resource allocation decisions. They may also meet the needs of citizens. However, the methodological challenges regarding particular situations and the preference based measures currently in use require further research in order to improve cost-benefit analysis and explore alternative preference elicitation methods for decision making. The comparative studies concerning the application of different techniques to obtain utility values show that techniques such as Discrete Choice Experiment are promising with regard to the use of the more traditional Visual Analogue Scale, the Time Trade-Off and the Standard Gamble. However, this technique still presents challenges and requires improvements. Further applications of this technique along with the

¹ Regional Health Authority, Coimbra

² Faculty of Economics, University of Coimbra

³ University of the Algarve-ESGHT

⁴ Centre for Health Studies & Research, University of Coimbra

E-mails: patriciaferreira@mail.telepac.pt, pedrof@fe.uc.pt, Inferrei@ualg.pt

implementation of hybrid methods are needed to support the decision making process. This paper provides a brief description of each of these techniques and illustrates some of their applications.

Keywords: Preference based decisions, Standard Gamble, Time Trade-Off, Discrete Choice Experiment, Economic Evaluation.

1. Introduction

In health policies as in other areas of the economy, decisions are based on the existence of limited resources. In any of the public, private or social components of a health system these decisions have clear consequences on the health of populations.

Economic evaluation is a process to prioritize and decide on efficient allocation of human, financial and technological resources. Its main purpose is to support the decision makers on the choice of efficient and equitable options, when comparing costs and consequences of alternative programmes and health interventions (Brazier *et al.*, 2007; Drummond *et al.*, 2005). The decision making process is rather complex as it is not limited to the decision on different alternatives according to the health gains originated. In this process multiple criteria is taken into account such as social equity, the quality of patient experience, impacts on the wider economy and the quality of evidence upon which to base a decision (Devlin & Sussex, 2011).

The process of choosing among technologies and among health interventions has contributed to a greater efficiency in the allocation of resources (Garrido *et al.*, 2008). The multiple challenges that public health systems embrace, such as the maximization of health, the reduction of inequalities, the current economic context and the substantial pressure on state budgets, assume particular complexity to the decision-makers regarding the choice of new technologies or health programs and on the appraisal of existing procedures. Baltussen and Niessen (2006) state that problems of this nature are handled with some difficulty by

the decision makers, using heuristic or intuitive approaches to reduce its complexity. Decision makers may also be influenced by individual motivations and political pressures that will constitute a disturbing factor. The risk of arbitrariness is therefore real. The variability of the decision in seemingly similar situations may sometimes be explained by contextual issues such as personal values, different risk perceptions, different priorities and variable budgets. For this reason, the choice of the most suitable method to decide on the technology, program or intervention in health is of the utmost importance, in particular with regard to the independence, transparency and impartiality of the definition and measurement of value.

The purpose of economic evaluation is to identify, measure, assess and compare costs and consequences when choosing various alternatives that lead to different use of resources (Brazier *et al.*, 2007). When looking for an informed decision, Drummond *et al.* (2005) emphasize the need to perform a systematic analysis for a clear identification of all the relevant alternatives and the importance of knowing the different perspectives of the stakeholders involved as well as determining the associated opportunity costs, comparing them with the consequences obtained in each alternative. Devlin and Sussex (2011) argue that in theory if it was possible to determine the cost-effectiveness of all health services and technologies, it would be possible to maximize the benefit of the population constrained to a specific budget. Regardless of the scenario involved it is necessary to determine the cost-effectiveness analysis with the greatest accuracy possible and accomplish the preferences of stakeholders. Given the methodological challenges in assessing particular situations (children, elderly, specific diseases) it is necessary to deepen the knowledge and improve the various existing techniques for the determination of cost-effectiveness as well as to explore alternative techniques of elicitation of preferences to support decision making. In this sense, it is intended to present and discuss the most common techniques of elicitation of preferences in health.

2. Preference based measures in health

The decision on allocation of health resources can be hampered if just considering the physiological impact of an intervention. This limited approach contradicts the comprehensive concept of health, adopted by the World Health Organization since 1946, which defines health as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (World Health Organization, 1946).

The Quality-Adjusted Life Year (QALY) is an indicator based on preferences for support decision making in public institutions when applying techniques of economic evaluation. This indicator allows us to optimise the available resources and may be used in different situations (e.g., in different diseases) and in individual and population scopes (Kobelt, 2013; Rowen *et al.*, 2014). QALY represents a real innovation in health economics and decision making process, when combining the quality of life with length of life (Brazier *et al.*, 2007). Quality of life is measured on a full health-dead 1-0 scale, where one equals full health and zero is equal to being dead, with negative values considered to be worse than death. In practice for any individual the prospect of living Y years in less than full health may be equated to a prospect of living X years in full health where $X < Y$ (Brazier *et al.*, 2007).

The instruments used for obtaining health state values have two components: a descriptive system for describing health or its impact on quality of life and an algorithm for assigning values to each state described by the system (Tsevat *et al.*, 1994; Brazier *et al.*, 2007). While the measurement of health status objectively assess physiological results or inquire the individual about their functional capacity and limitations, the measurement of value (value, preferences, utility) asks the individual in relation to the value he assigns to a specific health condition, assuming, for example, that individuals with similar limitations may score them distinctly. Drummond *et al.* (2005) explain that although the terms “value”, “preferences” and “utility” are sometimes used indistinctly, there are differences in their meaning: the term “preference” applies to the broader concept and the terms “value” and “utility” fall as types of preferences.

There are several instruments or generic preference based measures of health including the Quality of Well-Being or QWB, the 15-D, the Health Utilities Index or HUI, the EQ-5D and the SF-6D. The QWB, developed in late 1970, is the oldest of the QALY instruments (Kaplan *et al.*, 1976; Brazier *et al.*, 2007). This instrument has three multilevel dimensions relating to function (mobility, physical activity and social activity) and a list of symptom and problem complexes. Weights have been estimated from a sample of health states using a Visual Analogue Scale. The aim is to produce a value (and utilities) on a scale of preferences that ranges between zero and one, or death and full health (Drummond *et al.*, 2005; Seiber *et al.*, 2008). The 15-D is a generic instrument, consisting of 15 dimensions that cover most of the dimensions of a preference based measure. The valuation of 15-D has been assessed through the combined use of rating scales and magnitude estimation (Sintonen, 1994, 1995). The HUI is a family of generic health measurement instruments: HUI1, HUI2 and HUI3, being HUI3 the most widely used. The first one was developed for use in an economic evaluation of neonatal intensive care, the second one is now used as a generic based-preference for children and the third one is for adults (Horsman *et al.*, 2003; Brazier *et al.*, 2007). HUI3 has eight dimensions and has been valued using the Standard Gamble technique (Horsman *et al.*, 2003; Drummond *et al.*, 2005). The EQ-5D is an instrument developed by a group of researchers (EuroQol Group) for use in economic evaluation. This index derives from an initial descriptive system with six attributes (mobility, self-care, usual activities, social relationships, pain/discomfort and anxiety/depression), subsequently revised to include only five attributes (social relationships were excluded). Each attribute has three levels – no problems, some problems and extreme problems – that will enable the definition of 243 (=3⁵) possible health states. These are valued using a Visual Analogue Scale and the Time Trade-Off technique. The values will be situated on a scale that ranges between zero and one, or death and full health. Negative values may be assumed for health states considered worse than death (Drummond *et al.*, 2005; Ferreira, 2010; Ferreira *et al.*, 2013a). The Portuguese version of the EQ-5D was completed in 1998 and published in 2013, by Ferreira

and his collaborators (Ferreira *et al.*, 2013a). The SF-6D was developed from the SF-36 instrument and has six dimensions (physical functioning, role limitation, social functioning, pain, mental health and vitality), each with four to six levels, allowing the generation of 18.000 distinct health states (Brazier *et al.*, 2002). In any analysis, the use of the SF-6D cannot be made independently, and must be preceded by the implementation of the SF-36 questionnaire (Ferreira *et al.*, 2013b). The health states obtained with the SF-6D are valued using the Standard Gamble technique.

Figure 1 schematically represents the process of measuring value of a specific health condition.

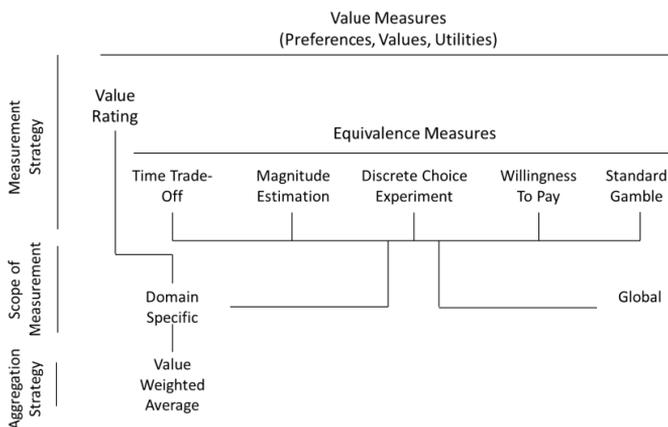


Figure 1: Ascertain a value for a health state⁵

Tsevat *et al.* (1994) point to the advantage of measuring value instead of measuring a health state being the former a method that best fits the commitments of the decision making process.

In 1944, von Neumann and Morgenstern during their research on Theory of Games developed a rational theory on decision making in a context of uncertainty, which became known as the expected utility theory or theory of von Neumann-Morgenstern utility. The normative model developed by these researchers determines the behavior of a rational

⁵ Adapted from Tsevat *et al.* (1994). Using Health-related Quality-of-life Information. *Journal of General Internal Medicine*, 9(10), 576–582, Figure 1.

individual in decision making, under a context of uncertainty. This model considers that decision agents have preferences that dominate a set of possible choices, satisfying the axioms of transitivity, independence and continuity (Drummond *et al.*, 2005).

Generally speaking, the term “utility” can be understood as the more an individual prefers a certain result the more utility will be associated with it. However, this concept reveals itself complex when it attempts a more specific definition and, especially, when looking for its measurement (Drummond *et al.*, 2005).

The methods used in the process of measuring preferences vary in several aspects: the type of preferences (value or utility), the context of certainty/uncertainty, and the response of individuals (scaling or choice). Schematically, the techniques for measuring preferences can be represented as follows (table 1).

Table 1: Methods of measuring preferences⁶

Response method	Question framing	
	Certainty (values)	Uncertainty (utilities)
Scaling	Rating scale	
	Category scaling	
	Visual analogue scale	
	Ratio scale	
Choice	Time trade-off	Standard gamble
	Paired comparison	Discrete choice
	Equivalence	
	Person trade-off	

According to Brazier *et al.* (2007) different techniques can generate distinct health values. Thus, it is important to reflect on the advantages and disadvantages of techniques more commonly chosen to integrate the instruments for measuring quality of life. The three major techniques for valuing health states are the Visual Analogue Scale (VAS), the Standard Gamble (SG) and the Time Trade-Off (TTO). More recently a new technique

⁶ Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 143, Table 6.1.

has been introduced: the Discrete Choice Experiment (DCE). However, which is the best technique?

The next section provides a brief description of each of these techniques.

3. Visual Analogue Scale

The VAS is a graduated ruler, usually drawn horizontally with 10 cm or vertically with 20 cm, with well defined end-points, on which respondents are able to indicate their judgement about a health state. The distances between intervals on a VAS represent the same difference and correspond to the difference in terms of perceived preference by individuals (Brazier *et al.*, 2007). The VAS has been used in economic evaluation for three decades and is the most common technique taking part in the QWB, HUI, 15-D and EQ-5D instruments for valuing health states. This technique can be used to assess chronic states considered better than death, states worse than death and temporary health states. In order to use VAS in economic evaluation it is necessary to ensure comparability among respondents. This requires clear and unambiguous end-points (full health, death) and a clear definition of the concept of full health in order to minimize the risk of different interpretation among respondents and even researchers (Brazier *et al.*, 2007).

To use this technique it is desirable to ensure that health states valuation can be placed on a zero to one scale, where zero is for states equivalent to death and one represents a state of full health. However, it is also necessary to allow for states valued worse than being dead. After obtaining a value for death and based on the interval properties of the scale, then all health states are calculated according to the following formula:

$$A_i = \frac{R_i - R(\text{dead})}{R(\text{full health}) - R(\text{dead})}$$

where A_i represents the adjusted VAS rating for health state h_i ; $R(\text{dead})$ is the raw rate given to being dead; R_i is the raw rate given to health state h_i and $R(\text{full health})$ represents the rate attributed to the

full health state. The value of A_i would lie between one (full health state) and zero (death) although it may assume negative values for states valued as worse than being dead (Brazier *et al.*, 2007).

Preferences for temporary health states may be measured using a VAS as long as respondents are fully acknowledged that the health states will last for a specific period of time (weeks, months or years; a period of time less than life expectancy) after which the person will return to full health (Brazier *et al.*, 2007).

4. Standard Gamble

The SG offers the respondent two alternatives, a certain intermediate outcome (state i) and the uncertainty of a gamble with two possible outcomes, one of which is better than the certain intermediate outcome - full health - and the other is worse - death (Torrance, 1986). This technique is used for measuring temporary health states, chronic health states and chronic health states better or worse than death. The method changes its format according to the situation and for a chronic health state preferred to death can be illustrated as shown in Figure 2.

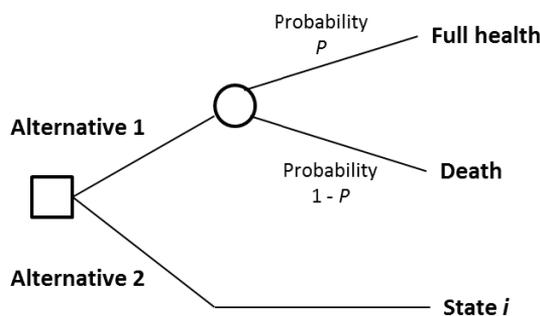


Figure 2: Illustration of a standard gamble ⁷

⁷ Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 150, Figure 6.2.

Alternative 1 offers two possible outcomes, with distinct probabilities: either the individual is returned to normal health and lives for an additional t years (probability P), or the individual dies immediately (probability $1-P$). Alternative 2 has the certain outcome of chronic state h_i for life (t years). In this game, the probability P of the best outcome is varied until the individual is indifferent between the certain intermediate outcome (alternative 2) and the gamble (alternative 1). In this case, the value of health state i (h_i) is equal to the P value measured on the utility scale, in which full health for t years corresponds to one and immediate death to zero (Ferreira, 2003; Drummond *et al.*, 2005). The technique used for the variation of value P is called ping-pong, since lower values are toggled with higher values until reaching indifference.

5. Time Trade-Off

The TTO is a less complex technique developed as an alternative to the SG. While the SG assumes a choice between a certain outcome and a gamble based on uncertainty, the TTO requires a simple choice between two certain alternatives. It is a less complex technique because respondents are not required to understand probabilities (Brazier *et al.*, 2007). The TTO is a technique that can be used in chronic health states considered better or worse than death and temporary health states (*mutatis mutandis* to the technique). Figure 3 illustrates the alternatives for a chronic health condition preferred to death.

For a chronic health state preferred to death, alternative 1 involves living in a less than full health state (h_i) for a period t (e.g., life expectancy of a particular chronic condition), followed by death. Alternative 2 involves full health for a period x where $x < t$, followed by death. In this technique, time x will be varied until the individual is indifferent between the two alternatives. In this case, the value of the health state i (h_i) is equal to x/t (Ferreira, 2003; Drummond *et al.*, 2005).

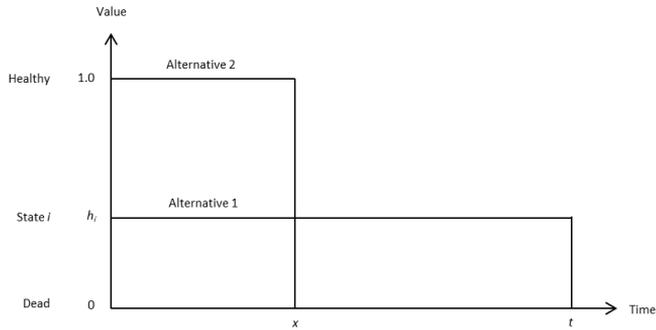


Figure 3: Time trade-off for a chronic health state preferred to death⁸

6. Limitations of the most common techniques

As mentioned before, VAS, SG and TTO are the standard techniques for eliciting preferences in health. However, there has been an intense debate about which is the best technique. Although the TTO and SG are recognizably superior techniques when compared to VAS, since they are based on choices that involve sacrifice, the values obtained from each of these techniques generate little consensus (Rowen *et al.*, 2014). Van Osch *et al.* (2004) reported that the utility values obtained by these techniques can be distorted by biases due to loss aversion (TTO, SG), scale compatibility (TTO), utility curvature for life duration (TTO) and probability weighting (SG). These concerns were also shared by other authors (Brazier *et al.*, 2012). Generally speaking, the cumulative effect of these limitations on obtaining utility values is not known when using TTO. It is also assumed that the utility values are higher when obtained through the SG. Even though props and visual aids have been developed to address literacy concerns, the TTO and SG have been criticized for their complexity towards the respondents and also because they exclude particular groups such as children, the elderly or different cultures (Brazier

⁸ Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 152, Figure 6.4.

et al., 2012; Rowen *et al.*, 2014). On the other hand, VAS is a technique that is not strictly based on preferences as it does not imply a choice between options (Brazier *et al.* 2007).

The limitations of these techniques have led to a growing interest in others such as the Magnitude Estimation, Person Trade-Off and Discrete Choice Experiment (DCE). Although these techniques are less frequent, DCE has been getting increasingly featured in studies when eliciting preferences in health.

7. The discrete choice experiment for elicitation of preferences

DCE is a method for eliciting preferences that has been used since 1960. It is based on mathematical psychology and was originally applied in the marketing area. Since then, DCE has been used for eliciting preferences of individuals in other academic areas such as transportation and environmental economics and more recently in health (Earl & Kemp, 1999; Kenny *et al.*, 2003; Lancsar & Louviere, 2008; Norman *et al.*, 2013). The use of this method of elicitation is particularly relevant because preferences play a key role in priority setting and resource allocation in health (Lancsar & Louviere, 2008). The use of DCE is also justified by its simplicity of implementation and for being a method easier to understand (Kjær, 2005).

Discrete choice occurs when an individual is faced with a choice between a finite set of alternatives, mutually exclusive and that contemplate all the possibilities. In DCE the respondents have to choose one alternative out of a given number of alternatives (two or more). DCE is usually implemented with questionnaires that describe the good/service by a number of attributes and according to the most relevant attributes, the respondent associates a utility.

The purpose of this method is to observe the response of the individual to the change of these attributes throughout levels to which they are assigned. Each combination, built with the various attributes and associated levels, will constitute a scenario (*i.e.* a different package of

good/service) whose ultimate aim is to allow the individual to feel able to make exchanges between scenarios according to what matters most to him/her.

Thus, it becomes possible to identify the level that influences the choice of each attribute, i.e., estimating the marginal rate of substitution of the attributes (Kjær, 2005; Lancsar & Louviere, 2008). The cost attribute (e.g., transportation costs, wages and taxes) plays an extremely important role in DCE as it allows to elicit preferences concerning the willingness to pay. This does not mean that the respondents are directly asked about their willingness to pay. Instead they are asked to carry out monetary exchanges for improvements in attributes (Kjær, 2005). Similarly, the life expectancy is also a way of measuring the value that is attributed to a health state.

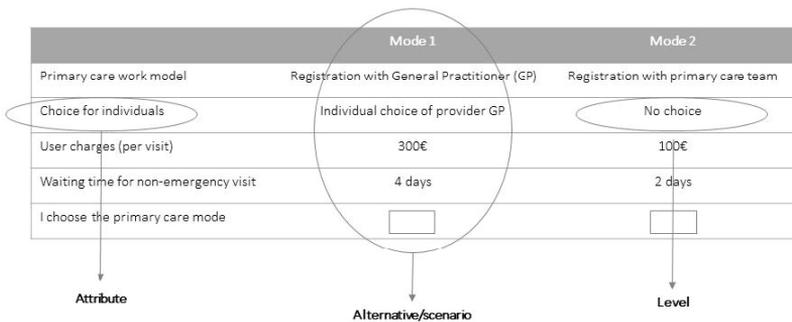


Figure 4: A choice set⁹

In order to analyse the results of DCE, each respondent choices converges into a single utility value, assigning weights to each attribute as a linear combination (Earl & Kemp, 1999; Kjær, 2005).

Choices made in the DCE are analysed using random utility theory, i.e., for an individual i conditioned on choice j , the utility can be decomposed into an explainable component V_{ij} and a non-explainable or random component ε_{ij} :

$$U_{ij} = V_{ij} + \varepsilon_{ij}, j=1, \dots, J$$

⁹ Adapted from Hjelmgren & Anell (2007). Population preferences and choice of primary care models: A discrete choice experiment in Sweden. *Health Policy*, 83, 314–322.

The non-explainable component ε_{ij} may be due to unobservable or unobserved attributes, unobserved variations in preferences, specification error and measurement error.

It is assumed that an individual chooses a particular option 1 if, and only if, its utility is higher than the utility of any other option on the set of J alternatives. Whereas Y_i is a random variable that denotes the choice outcome and assuming a joint probability distribution for ε_i , the probability P that utility is maximized by choosing option 1 is given by:

$$\begin{aligned} P(Y_i=1) &= P(U_{i1} > U_{ij}) \\ &= P(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij}) \\ &= P(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}), \forall j \neq 1 \end{aligned}$$

To move from a probabilistic choice model to an econometric choice model, the observable or systematic component of utility must be specified:

$$V_{ij} = X'_{ij}\beta$$

where X'_{ij} is a vector of variables representing observed attributes of option j and β is a vector of preferences parameters to be estimated (McIntosh *et al.*, 2010). V_{ij} can also depend on the characteristics of individual decision makers (Z'_i – vector of characteristics of individual), interacted with attributes of the good or service (X'_{ij}) but this would be ignored to simplify the discussion:

$$V_{ij} = X'_{ij}\beta + Z'_i\gamma$$

It is not wise to apply DCE without fully understanding the method. Therefore, it is important to study more deeply the theory, the design of the method and the results that can be obtained in order to build a valid model for research studies (Lancsar & Louviere, 2008).

Each design of DCE takes into account the specific context of the ongoing study and divides the several distinct phases in order to simplify the use of the method. Several authors argue for similar designs yet with slight differences. However, there is some consensus to split the design of DCE in five main stages: i) identification of attributes; ii) identification of levels; iii) experimental design; iv) data collection; and v) analysis of data (Kjær, 2005; Lancsar & Louviere, 2008). For the successful implementation of the method it is crucial that researchers have a broad notion of what is being studied and that the respondents

are fully aware of the good/service under review (Lancsar & Louviere, 2006).

The first stage of the design involves defining the attributes of interest. These must be comprehensive and measurable, and can be qualitative or quantitative. For the identification of the most appropriate attributes in the characterization of a good/service, there are several possible sources of information including the existing literature, focus groups, interviews with key persons (e.g., decision makers) and experts (Kjær, 2005; Lancsar & Louviere, 2008). The main objective in the definition and identification of attributes is that they are relevant to the decision makers and that, simultaneously, are meaningful for the majority of the respondents. There is no maximum limit on the number of attributes but some researchers, particularly in the area of health economics, suggest not to exceed twelve attributes (Lancsar & Louviere, 2008). In identifying the attributes there are several aspects to be taken into consideration in particular if these are irrelevant or not, i.e., to what extent their exclusion/inclusion interferes with the results obtained and if the attributes are, among themselves, mutually dependent or causality related. These characteristics are particularly important because they may influence individuals by inducing a decrease in utility or even the transfer of utility between attributes (Kjær, 2005).

The second stage involves assigning levels to the identified attributes. Thus, for each of the attributes identified in the previous step reasonable and relevant levels are assigned. The range of these levels should be such as to enable the respondents to carry out exchanges between the various combinations of attributes. The number of levels of an attribute influence the significance of the attribute (considering equal variation intervals). However, the greater the number and levels of attributes the more complex will be the design (Kjær, 2005; Lancsar & Louviere, 2008).

With regard to obtaining values for utility measuring instruments already validated, these two phases will be naturally overcome.

The experimental design stage involves the design of the hypothetical choice sets including the formation and pairing alternatives. The main

objective is to create the DCE so that the number of alternatives is as small as possible without questioning the ability of the respondent to infer utility on all the alternatives presented (Kjær, 2005). When pairing the alternatives it is also important that the differences in attribute levels for each choice set are not multi-correlated. The design will allow the estimation of a matrix, resulting from all possible combinations of levels of the different attributes of alternatives (choice set). The number of alternatives significantly increases as the number of attributes and the number of associated levels increase:

$$a=l^b$$

where a represents the number of alternatives; l is the number of levels and b represents the number of attributes (Lancsar & Louviere, 2008). Full factorial design refers to a design in which all the possible alternatives are represented. However, the full factorial design can only be easily applied to small experiments, with a very limited number of attributes and levels. In studies involving a larger number of attributes and levels it is necessary to reduce the size of the design. In this case the fractional factorial design is the most suitable option. Fractional factorial design involves a selection or a subset maintaining the properties of the original full factorial design. Although some loss of statistical information is involved, the fractional factorial design allows the estimation of effects of interest as efficiently as possible.

In order to ensure the optimisation of the design four principles need to be simultaneously considered: level balance (the levels of an attribute occur with equal frequency in the design); orthogonality (the difference in the levels of each attribute varies independently over choice sets); minimal overlap (a level does not repeat itself in a choice set); and utility balance (the utilities of the alternatives within each choice set are approximately equal). Full factorial design and fractional factorial design can be obtained through manual classifications or preferably using specific software (Kjær, 2005; Lancsar & Louviere, 2008).

The next stage involves the collection of data. Studies on the design of DCE recommend the inclusion in the model of an additional alternative on the choice set, particularly in circumstances where it is realistic

to consider that the good/service is not consumed. This alternative is considered a no choice and prevents the individual from being forced to choose something that is not important to him/her (Kjær, 2005). With the inclusion of this alternative the respondent may “opt out”, choose “no”, “not to participate” or keep the existing situation (*status quo*). In some complex models the limitation of this choice resides in the fact that the individual can choose the easiest answer, i.e., can choose to apply a heuristic to avoid making a choice deemed difficult. Another perspective on the inclusion of an additional alternative is the possibility that this reflects a situation of indifference to the individual (Kjær, 2005; Lancsar & Louviere, 2008). For the success of data collection it is essential to ensure that the respondent understands what is expected of him/her which requires a thorough explanation on the part of researchers with regard to the context and objectives of the study, the description of the attributes and the contextualization of each of the scenarios (choice sets). At this stage, researchers may also consider important to perform a validation using rationality tests. Therefore, they need to include additional choice sets to test the various axioms, such as completeness, transitivity and continuity (Lancsar & Louviere, 2006). Finally, it is also important to include questions regarding socio-demographic data (*e.g.*, age, gender, education, income level, occupation). The collection of data can be done using interviews, questionnaires, or a combination of these methods, and using the Internet.

With regard to the sample, it should reflect the population for which the results will be generalized and to whom the opportunity costs matter. The sample size will depend on the number of questions that each individual responds, the size of the population, the response rate that is expected to obtain in the study and the statistical power intended to obtain from the model (Kjær, 2005; Lancsar & Louviere, 2008).

The last stage involves data analysis and includes the choice of the probability model and the interpretation of data. Several models are available for the estimation of DCE whose choice depends, among other factors, on the design chosen for the method (Kjær, 2005). If the DCE happens to be dichotomous (yes or no answer for a given alternative) or present

a choice set consisting of two alternatives, then binomial discrete choice models (logit or probit) are used. When the choice set is made up of three or more alternatives then it is suitable to use multinomial (logit or probit), mixed logit, nested logit, and heteroscedastic extreme-value models. By increasing the number of alternatives in the choice set, these models put a series of challenges that do not exist in the binary choice (Kjær, 2005; Ryan *et al.*, 2008). Generally, the estimation of discrete choice models is based on the maximum likelihood method, although other methods may also be considered (Kjær, 2005). While DCE values are estimated on an unobserved anchored latent scale, they can be anchored on the health utility scale (zero to one scale) by incorporating duration (time) as an attribute. This method is referred to DCE_{TTO} (Bansback *et al.*, 2014).

8. Discussion

There are many concerns regarding the standard methods used to provide utilities. First, different utility values are obtained with SG and TTO and this raises the question of which method to use. Also, when a respondent has experienced a certain outcome (*e.g.*, disabled person, person with a chronic disease or a surviving cancer patient), he/she will approach the valuation task differently than a person who has not experienced the consequences of a disease or treatment. It is not unusual for a patient to refuse to sacrifice life-expectancy in order to be relieved of his/her health problem. This may be due to the time frame being longer than the time he/she expects to live. The use of a valuation task that offers an alternative like “full health” to a person suffering from a disease also raises main ethical concerns. This would not be an issue if the task was to be applied for the general population. Nevertheless, the potential limitations of standard utility measurement techniques in particular groups need to be carefully considered.

Although DCE is not free from bias either, this problem can be minimized by the experiment design. Another main advantage on the use of DCE is the ability to control the experiment. The possibility of con-

structuring a DCE offers the opportunity to focus to the precise issue of interest, like a group of persons suffering from a specific disease. This also allows for a range of attributes and levels to be valued within the same survey and therefore provides a richer set of information. Several other benefits are identified in this technique: ease of understanding (mainly by respondents), wider application (great number of people) and the fact that preferences are not influenced by time. Several studies have also begun to use DCE in particular groups, such as children and the elderly (Rowen *et al.*, 2014). Still, there are many differences between these valuation techniques and Table 2 resumes the main ones.

Table 2: Comparison between VAS, TTO, SG and DCE

Technique	Summary of differences
VAS	<ul style="list-style-type: none"> • Choiceless method
TTO	<ul style="list-style-type: none"> • Utilities biased upwards by loss aversion and scale compatibility • Tasks involved are too complex for certain populations • Exclude particular groups • Typically conducted with an interviewer (more time and cost consuming)
SG	<ul style="list-style-type: none"> • Utilities biased upwards by probability weighting and loss aversion. • SG utilities are systematically higher than TTO utilities • Tasks involved are too complex for certain populations • Exclude particular groups • Typically conducted with an interviewer (more time and cost consuming)
DCE	<ul style="list-style-type: none"> • Allow a range of attributes and levels to be valued within the same survey (provide a richer set of information) • Task cognitively simple • Typically conducted without an interviewer (less time and cost consuming) • Wider application

The comparison evidences that techniques such as DCE are promising with regard to the use of the TTO and SG (Braziet *et al.*, 2012; Krabbe *et al.*, 2014). But, apart from all the benefits identified DCE still presents challenges regarding the anchorage on the health utility scale with one for full health and zero for dead, which require further analysis (Rowen *et al.*, 2014).

The future will pass by carrying out additional studies using this technique and for assessing the implementation of hybrid methods (Bansback *et al.*, 2014).

9. Conclusion

In a context of scarce resources and particularly in health, the decision making process has important consequences on individuals. Whether in choosing new technologies or health programs, whether in the re-evaluation of existing procedures, the decision making process is a real challenge under the current economic context. The concern about the use of rational, exempt, replicable and transparent approaches clearly stands in the decision making process. On the other hand, to meet the concerns of citizens it is important to make a rigorous use of preference based measurement instruments in health and economic appraisal in particular. The review of currently available techniques and the conduction of additional studies with more promising methods are essential in supporting the decision making process.

References

- Baltussen, R. & Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation*, 4, 14.
- Bansback, N., Hole, A. R., Mulhern, B. & Tsuchiya, A. (2014). Testing a discrete choice experiment including duration to value health states for large descriptive systems: addressing design and sampling issues. *Social Science & Medicine*, 114, 38-48
- Brazier, J., Roberts, J., Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21(2), 271-292.
- Brazier, J., Ratcliffe, J., Salomon, J. A. & Tsuchiya, A. (2007). *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford: Oxford University Press.
- Brazier, J., Rowen, D., Yang, Y. & Tsuchiya, A. (2012). Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. *European Journal of Health Economics*, 21(2), 271-292.
- Devlin, N. J. & Sussex, J. (2011). *Incorporating Multiple Criteria in HTA: Methods and Processes*. London: Office of Health Economics.
- Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J. & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- Earl, P. E. & Kemp, S. (1999). *The Elgar companion to consumer research and economic psychology*. Cheltenham; Northampton, MA, USA: Edward Elgar.
- Ferreira, L. (2003). [Utilities, QALYs and quality of life measurement]. *Revista Portuguesa de Saúde Pública*, 3, 51-63.

- Ferreira, P. (2010). [Measurement of health status and quality of life. In: 30 years of the National Health Service: a commented path] (461–484). Coimbra: Almedina.
- Ferreira, P. L., Ferreira, L. N., Pereira, L. N. (2013a). [Contribution for the validation of the Portuguese version of EQ-5D]. *Acta Médica Portuguesa*, 26(6), 664–676.
- Ferreira, L. N., Ferreira, P. L., Pereira, L. N., Rowen, D., Brazier, J. E. (2013b). Exploring the consistency of the SF-6D. *Value in Health*, 16(6), 1023–31.
- Garrido, M. V., Kristensen, F. B., Nielsen, C. P. & Busse, R. (2008). *Health technology assessment and health policy-making in Europe*. Copenhagen: WHO European Observatory on Health Systems and Politics.
- Hjelmgren, J. & Anell, A. (2007). Population preferences and choice of primary care models: A discrete choice experiment in Sweden. *Health Policy*, 83, 314–322.
- Horsman, J., Furlong, W., Feeny, D., Torrance, G. (2003). The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health and Quality of Life Outcomes*, 1, 54.
- Kaplan, R. M., Bush, J. W., Berry, C. C. (1976). Health status: types of validity and the index of well-being. *Health Services Research*, 11(4), 478–507.
- Kenny, P., Hall, J., Viney, R. & Haas, M. (2003). Do participants understand a stated preference health survey? A qualitative approach to assessing validity. *International Journal of Technology Assessment in Health Care*, 19(4), 664–81.
- Kjær, T. (2005). A review of the discrete choice experiment - with emphasis on its application in health care. *Health Economics Papers*, 1.
- Kobelt, G. (2013). *Health Economics: An Introduction to Economic Evaluation* (3rd ed.). London: Office of Health Economics.
- Krabbe, P. F. M., Devlin, N. J., Stolk, E. A., Shah, K. K., Oppe, M., Van Hout, B., ... Xie, F. (2014). Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-states values. *Medical Care*, 52(11), 935-943
- Lancsar, E. & Louviere, J. (2006). Deleting “irrational” responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics*, 15(8), 797–811.
- Lancsar, E. & Louviere, J. (2008). Conducting Discrete Choice Experiments to Inform Healthcare Decision Making A User’s Guide. *Pharmacoeconomics*, 26(8), 661–677.
- McIntosh, E., Clarke, P., Frew, E. & Louviere, J. (2010). *Applied Methods of Cost-Benefit Analysis in Health Care*. Oxford: Oxford University Press.
- Norman, R., Viney, R., Brazier, J., Burgess, L., Cronin, P., King, M. & Street, D. (2013). Valuing SF-6D Health States Using a Discrete Choice Experiment. *Medical Decision Making*. [published online before print] doi:10.1177/0272989X13503499
- Rowen, D., Brazier, J. & Van Hout, B. (2014). A comparison of methods for converting DCE values onto the full health-dead QALY scale. *Medical Decision Making*. [published online before print] doi: 10.1177/0272989X14559542
- Ryan, M., Gerard, K., Amaya-Amaya, M. (2008). *Using discrete choice experiments to value health and health care*. Dordrecht: Springer.
- Seiber, W. J., Groessl, E. J., David, K. M., Ganiats, T. G. & Kaplan, R. M. (2008). *Quality of Well Being Self-Administered (QWB-SA) Scale User’s Manual*. San Diego.
- Sintonen, H. (1994). *The 15-D Measure of Health Related Quality of Life : Reliability, Validity and Sensitivity of its Health State Descriptive System*. Melbourne: Centre for Health Program Evaluation.

Série Investigação

•

Imprensa da Universidade de Coimbra

Coimbra University Press

2015

