INCORPORATING DISCRETE CHOICE EXPERIMENTS INTO AN ECONOMIC EVALUATION: A CASE STUDY OF PREFERENCES FOR ALTERNATIVE MONITORING SERVICES FOR INDIVIDUALS WITH OCULAR HYPERTENSION

Rodolfo Hernández[1,2], Mandy Ryan[2], Luke Vale[3], Paola Botello Pinzon[4], Jennifer M Burr[5].

1 Health Services Research Unit, University of Aberdeen, UK ; 2 Health Economics Research Unit, University of Aberdeen, UK; 3 University of Newcastle, UK; 4 NHS Grampian, UK; 5 University of St Andrews, UK.

Correspondence to: Rodolfo A. Hernández, r.a.hernandez@abdn.ac.uk

Abstract

Whilst Discrete Choice Experiments (DCEs) have been increasingly used in Health Economics, their incorporation within an Economic Evaluation framework is limited. This paper reports on a DCE that assessed general public preferences for monitoring schemes for individuals with ocular hypertension who are at risk of developing glaucoma. The DCE was part of a larger economic evaluation. Issues raised when incorporating a DCE into an economic evaluation models are discussed. In addition, methodological work exploring the number of choices to include in a DCE is presented and it is argued that health economists may be being too cautious in this respect.

Background

In a recent review of the application of discrete choice experiments (DCEs) in health economics it was concluded that whilst DCEs are being increasingly used in health economics to elicit preferences, their integration into economic evaluations is limited (de Bekker-Grob *et al.* (2012). Vale (2005) incorporated an existing DCE into a Markov Model to compare alternative methods of surgery for inguinal hernia. However, this DCE was not originally designed for incorporation into the hernia economic model, imposing restrictions on the model structure that was compared with conventional cost-utility analysis. Regier (2008) applied a DCE to estimate benefits for different diagnoses in a simple decision analytic model comparing diagnostic interventions (Regier 2008). A decision analytic modelling approach, while appropriate to the context, is rarely suitable to capture the complexity of decision problems more commonly faced. This paper reports the results of a DCE that incorporated monetary values generated into a discrete event simulation economic evaluation modelling framework. The application is general public preferences for monitoring schemes for individuals with ocular hypertension who are at risk of developing glaucoma (Burr, Botello-Pinzon et al. 2012).[1] In this DCE consideration was also given to the

[1] Glaucoma is a chronic progressive optic neuropathy leading to impaired vision and blindness if inadequately treated. A raised intraocular pressure (IOP) is the only modifiable risk factor affecting around one million

number of choices to include. In the next section the methods are described. The results are then presented. Points for discussion around the number of choices to present in a DCE and incorporation of results within an economic modelling framework are discussed. This paper is work in progress and all comments are gratefully received.

Methods

Use of qualitative data to define attributes and inform questionnaire development

This study used a mixed methods approach to derive attributes and levels, integrating findings from an advisory panel (Box 1) and a focus group (Box 2) with relevant glaucoma health outcome attributes from an existing glaucoma utility measure (Burr, Kilonzo et al. 2007), an existing economic model (Hernandez, Burr et al. 2008) and expert opinion within the study team.

The topic guide developed from the advisory group formed the basis of the focus group discussion. The focus group discussion was audio-taped, transcribed and analysed independently by two researchers (APB, RH), using a modified framework approach (Ritchie, Spencer 1994). Using this approach, after initial familiarisation with the transcripts, the data was grouped according to common themes, reflecting the main research questions (related to attributes of a surveillance programme for ocular hypertension) and key issues that transpired from the discussion.

Box 1: The Advisory Panel

The advisory panel identified potential attributes (characteristics influencing preferences for alternative monitoring services) and this information was used to develop the framework (topic guide) for a focus group held with service users. Advisory panel participants were purposively selected to provide a broad UK perspective on managing ocular hypertension but who were not otherwise involved in the project. The panel included: two service users, one optometrist, one non-clinical health service manager, one specialist glaucoma nurse, one ophthalmologist, and one community optometrist. Six members of the local study team attended the advisory panel meeting which was recorded and transcribed. Minutes of the meeting with key points of agreement were circulated to participants following the meeting. Discussion was driven by the participants on the desired elements of a monitoring service with the members of the study team being involved as observers or providing point of clarification. Key themes (attributes) are summarised in Table 1.

people in the UK. Once ocular hypertension is diagnosed, long term surveillance to identify the need for treatment or to monitor treatment effectiveness is required.

Box 2: The Focus Group

Participants of the focus group were all users of a service identified from the Grampian Glaucoma Referral and Monitoring Scheme Service (a service started in June 2004 to tackle unnecessary referrals to a hospital glaucoma clinic). Both men and women with ocular hypertension (who may or may not be undergoing treatment) or suspected of ocular hypertension, of different ages and geographical locations within the Grampian area were purposely sampled. Participants were approached by post with an invitation letter from the consultant ophthalmologist and an information leaflet explaining the purpose of the study. In total, 93 people were invited and 11 agreed to participate in the study. Of these 11, six were able to attend the suggested date and time for a focus group. In the event five attended on the specified time and date (3 men/ 2 women, mean age 61 years old, range 54 to 67).

Discussion was facilitated and moderated by a clinically qualified health services researcher (APB), an economist (RH) and an ophthalmologist. Participants were encouraged to discuss their views and express their preferences about different characteristics of a monitoring programme. Throughout the discussion, the facilitators dealt points of clarification as requested by participants.

Table 1 Summary of attributes the advisory panel felt were important in a monitoring service

Attribute	***Discussion***
Risk	Loss of vision causing, for example, loss of the ability to drive.
Trust/confidence	Trust/confidence in who was seen. Patients are influenced by advice from their optometrist. Continuity of care would be preferable and this would contribute to building up a personal relationship and, therefore, an increase in trust.
Time/cost of test	Some people might be more concerned about time commitments rather than cost. However, people value their sight, and people may be willing to give their time and money to maintain their sight.
Place of testing	People would prefer not to have to go to a hospital for appointments because of parking problems and it would be easier if the location was on a bus route. However, people will go wherever they are told is best. Whilst location was important, overall confidence in a person was felt to be more important.
Type of appointment	The merits of a fixed monitoring appointment arranged by the monitoring service rather than patient arranged appointments were discussed. It was felt that a person's availability and flexibility would be affected by whether they were employed or not.
Frequency of appointment	The number of visits was felt to be of lesser importance than serious consequences (e.g. visual loss) that might result if not seen often enough. Differences in recommendations for community care for 'low risk' patients were noted between England (one visit per year) and Scotland (2-3 times per year).
Length of appointment	Length of appointment was felt to be important. People might prefer to see a specialist nurse if it involved a shorter waiting time.

Attribute	*Discussion*
Type of test	Type of test was felt to have an influence on the overall appointment. From a patient perspective, it was noted that one might undergo any recommended test to avoid going blind.

The final attributes and levels reported in Table 2. For more details on the qualitative work see Burr et al. (2012). Attributes are a combination of health outcomes and patient experience factors, allowing a holistic measure of value to be estimated.

Table 2 Summary of the attributes and levels of the monitoring service

Attribute included in the DCE	**Value for each level included in the DCE**	**No monitoring alternative**
Health Outcomes		
10-year risk of developing(*n*umber of people out of 10,000):		
Glaucoma	740; 960; 1190; 1410	1600
Severe glaucoma	25; 60; 100; 130	180
Visual impairment	2; 6; 10; 15	25
Unwanted effects of treatment	***None*** **–**means that you have not noticed any discomfort or difficulties. ***Some*** **–**means that you have noticed occasional discomfort or difficulty. ***Quite a lot*** **–**means that you are aware of these discomforts or difficulties most of the time. ***Severe*** **–**means that you need or think you need additional treatment to control one or more of these difficulties.	***None***
Patient Experience Factors		
Communication and understanding of information provided by the health professional	Made me feel **at ease** and made sure I **understood** the purpose of monitoring. Made me feel **at ease** but did **not** make sure I **understood** the purpose of monitoring. Did **not** make me feel **at ease** but made sure I **understood** the purpose of monitoring. Did **not** make me feel **at ease** and did **not** make sure I **understood** the purpose of monitoring.	Not applicable
Location	Hospital eye clinic Local optician	No testing
Price Proxy		
Cost per year *	£15; £30; £50; £70	No cost

Health outcomes

The chance of developing glaucoma, particularly of losing vision, was identified as important. To explore preferences for the health effects in the DCE we built on our previous work with glaucoma patients which developed the Glaucoma Utility Index (Burr, Kilonzo et al. 2007). The aim here was to include attributes reflecting a patients' perspective of the health impact of glaucoma and its treatment (Box 3).

Box 3: Risk and effects of treatment

The risk of developing glaucoma was redefined as three attributes (risk of developing: glaucoma; severe glaucoma; and visual impairment) with their levels determined using an existing economic model, tailored for individuals with OHT (Hernandez, Burr et al. 2008). Full details of these can be accessed elsewhere (Burr, Botello-Pinzon et al. 2012). The remaining two dimensions in the Glaucoma Utility Instrument, describing local and systemic side effects of treatment for OHT and glaucoma, were combined as one attribute (unwanted effects of treatment) to capture the unwanted consequences of treatment within the DCE. Levels were defined from the glaucoma specific utility measure (Burr, Kilonzo et al. 2007)

Using aids to communicate risk

Communicating risk to participants is known to be challenging (Edwards, Elwyn 2001, Lloyd 2001) and within DCEs de Bekker-Grob *et al*. (2012) argued that little progress has been made in explaining risk to respondents in DCEs. We explored the use of visual aids to help explain the concept of risk in the focus group. Firstly pictorial cards were used to describe the different stages of progression of ocular hypertension (Figure 1) and, following Ancker et al. (2006) (Ancker, Senathirajah et al. 2006), a graphical presentation of the probability of developing each disease stage (Figure 2). Moreover, we attempted to use plain language and frequencies (with the same base) to present the risk information (Fagerlin, Zikmund-Fisher et al. 2011).

The pictorial cards were opportunistically piloted amongst six members of the Institute of Applied Health Sciences, University of Aberdeen. The final questionnaire was pre-piloted opportunistically among 10 members of staff from the University of Aberdeen. Feedback suggested that the pictorial representation of disease stages (Figure 1) and graphical risk explanation (Figure 2) were very useful in understanding the context.

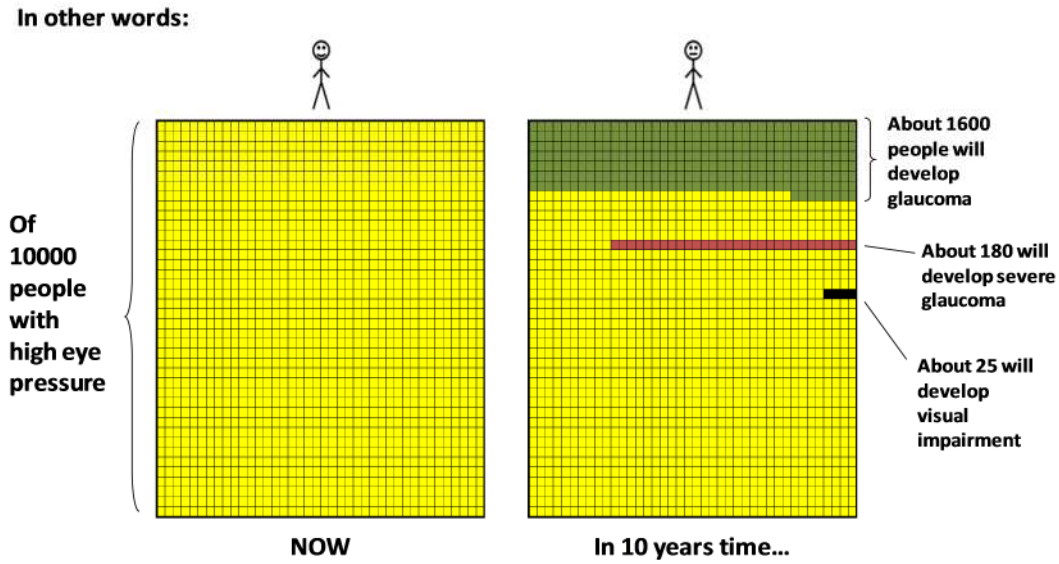
Figure 1 Describing the clinical condition

High Eye Pressure: This is a condition also known as Ocular Hypertension, where the eye pressure is higher than it should be. Ocular Hypertension is usually discovered during a routine eye test. Monitoring is needed to make sure glaucoma is not developing and to give treatment if necessary. Monitoring involves measuring pressure, usually a check of the vision to the sides, and an assessment of the nerve at the back of the eye with a bright light.	**No effect on vision to the sides** This is a picture showing that a person with high eye pressure (Ocular Hypertension) has normal vision and can see everything in this busy street 
Glaucoma: Glaucoma is an eye disease where there is damage to the nerve at the back of the eye, leading to a reduced vision to the sides. Once glaucoma is diagnosed, treatment will be required. This is usually in the form of daily eye drops. Eye surgery may be needed.  **About 1600 out of 10000 people with ocular hypertension will develop glaucoma in 10 years if not treated.**	**Glaucoma** Some effect on the vision to the sides. Note that in this picture, the red car on the right and the children cannot be seen. 

Figure 1 Describing the clinical condition (cont)

Severe glaucoma: A condition where there is a marked restriction of the vision to the sides due to glaucoma, leading to some difficulties with daily activities. Driving is not allowed (if glaucoma is severe in both eyes) because eyesight would not meet the required standards.  **About 180 out of 10000 people with ocular hypertension will develop severe glaucoma in 10 years if not treated.**	**Severe glaucoma** Following from the previous picture, there is more loss of vision to the sides. 
Visual impairment: This is an advanced stage of glaucoma where one can still see to read as the central vision is not affected very much but a lot of vision to the sides is lost. The main problem is difficulty getting around without assistance because the ability to see things at the sides is poor.  **About 25 out of 10000 people with ocular hypertension will develop visual impairment in 10 years if not treated.**	**Visual impairment** This is a more advanced stage of your condition, central vision remains but vision to the sides is lost. 

Figure 2 Graphical explanation of risk of developing glaucoma, severe glaucoma and visual impairment given to DCE participants



Patient Experience Factors

Communication and understanding of information provided by the health professional and *location* of screening were identified as important in the qualitative work. Whilst this DCE was being developed one of the authors (MR) was involved in a MRC funded project concerned with valuing patient experiences. A key finding from this work was that whilst there is an extensive literature valuing process-type attributes using DCEs, most studies value descriptors of healthcare provision. Few DCEs have defined attributes from the perspective of individuals, resulting in little consideration to how the attribute enables individuals to feel. In this study we aimed to define the '*communications and understanding of information*' attribute from the perspective of the individual, and how improvements would make them feel. Thus, the attribute was defined in terms of both whether the individual understood the information provided and whether they felt at ease. Location was defined as hospital eye clinic or local hospital (as was realistic).

Inclusion of price proxy

A price proxy, cost per year of the surveillance programme, was included so that monetary values could be generated and incorporated into a cost-benefit analysis (CBA). To identify levels for the cost attribute a direct willingness to pay question was included at the focus group stage.

Experimental design

A main-effects D-efficient design using SAS software (Kuhfeld 2010) was developed, ensuring uncertainty around parameter estimates was minimised (by minimising the determinant of the covariance matrix) (Kuhfeld, Tobias et al. 2010). This resulted in 32 choice sets for the pilot DCE. Zero beta parameter values were initially assumed (i.e. assuming the attribute level was not a determinant of preferences). The results from the pilot work (N=184; multinomial logit model) informed parameters for the main study experimental design, where 32 choices were again generated. It has been argued that incorporating the pilot data into the statistical design of the DCE improved D-efficiency (a measure of the statistical efficiency of the design) (Bliemer, Rose et al. 2009, Rose, Bliemer et al. 2008, Rose, Bliemer 2008) and in our case study the D-efficiency increased from 2.48 for the pilot design to 2.67 for the design subsequently used for the main survey.

For each of the 32 choices individuals were asked to choose between two monitoring scenarios and an opt-out 'no monitoring' alternative. This opt-out was added to each choice set to allow individuals to choose the realistic option of not being monitored. An example of a choice set is shown in Figure 3. The order of the choice questions presented to respondents was randomly generated.

Figure 3 Example of a DCE question.

Which monitoring service would you choose? (please tick one box below)

	Monitoring Service A	**Monitoring Service B**	**No monitoring Service**
Number of people **out of 10000** developing glaucoma in **10 years**	740	1410	1600
Number of people **out of 10000** developing severe glaucoma in **10 years**	25	130	180
Number of people **out of 10000** developing visual impairment in **10 years**	2	15	25
Unwanted effects of treatment	None	Severe	None
Communication and **understanding** of information provided by the **health professional**	Made me feel **at ease** and made sure I **understood** the purpose of monitoring	Did **not** make me feel **at ease** and did **not** make sure I **understood** the purpose of monitoring	Not applicable
Place of testing	Hospital eye clinic	Local optician	No testing
Cost per year	£15/year	£30/year	No cost
	Service A	**Service B**	**No Monitoring**
(tick one box only)	√		

Final questionnaire, sample and setting

Socio-demographic data (gender, age, region of UK, level of education, socio-economic status), experience of eye testing and health status using the EQ-5D 3L (www.euroqol.org) were collected. The final version of the questionnaire is available from the authors.

Given poor response rates from unsolicited general population surveys (de Bekker-Grob, Ryan et al. 2012) a web-based online survey method was used to collect the data, using a market research company (http://www.researchnow.co.uk/). The company's internal protocol was used to stratify by age and gender to obtain a UK representative sample of 800 individuals for the main survey, providing sufficient sample size to conduct sub-group analyses (e.g. 90 respondents per pre-specified sub-group (Pearmain 1991)).

How many choices to present individuals?

When developing a DCE issues are raised concerning how many choices an individual can respond to before coming bored and fatigued. There may also be some learning process, such that as the respondent progresses through the experiment they are better able to respond to such choices. The evidence on the optimal number of choices to include in a DCE is limited (Stopher, Hensher 2000). While transport economists seem to select designs with the fewest possible choice sets, in marketing research it has been suggested that respondents can deal with hundreds of choices (Louviere, 2004). Within the marketing, little evidence of fatigue has been found Hensher et al (2001) (Hensher, Stopher et al. 2001) whilst Hanley et al (2002) (Hanley, Wright et al. 2002) obtained weak evidence that the number of choice sets had a statistically significant effect on preferences when exploring the effects of choice complexity within environmental economics. An inverted u-shaped pattern was found for the error variance when plotted against the number of choice sets by Causade et al (2005) (Causade, Ortuza et al. 2005) (with an optimum around 9-10 choices); however, its effect on the individual's ability to choose was smaller than other design dimensions like the number of attributes and levels.

Within health economics there has been a slight shift towards designs using a larger number of choices, with studies using more than 16 choices per respondent rising from 6% (2 studies) for the period 1990-2000 to 18% (21 studies) for 2001-2008 {{2253 de Bekker-Grob,Esther W. 2012}}. However, there has been very little work looking at the optimal number of choices in health economics. Coast et al (2006) conducted a randomised controlled trial to determine whether response rates and/or results differ between questionnaires containing different numbers of choices: a short version capable of estimating main effects

only (8 choices) and a longer version capable of estimating two-way interactions (16 choices). The authors found no evidence of a difference in response rates with both questionnaires giving similar preference inferences (response rate difference 1.9%, 95% CI: -7.3 to 11.2; p = 0.68). However, it should be noted that the long questionnaire had fewer options than our DCE.

To explore this, in addition to the 32 choices derived from the experimental design, four rationality tests were added. The rationality tests applied Sen's contraction test (Sen 1993, Bech, Kjaer et al. 2011). The application of these tests is limited when mailed surveys are used to collect data since the follow-up question should be related to the initial response. Failure to do this will mean that some responses cannot be used to test rationality of some respondents (de Bekker-Grob, Ryan et al. 2012). However, the use of an internet based survey allowed individual specific choices. Individuals were initially presented with a choice set involving three choices. This choice was then repeated later in the questionnaire, with the choice set contracted to 2 options. The 2 options presented in the contracted choice set depended on what respondents chose initially, with the option they chose and one of the other options from the original three choices being presented. The individual should choose the same option when the choice set is contracted. The tests were included at four points within the questionnaire (Choices 2 and 5; Choices 11 and 15; Choices 20 and 26 and Choices 28 and 34) and consideration given to response patterns of correct/failed responses.

Analysis – Incorporating WTP into a Discrete Event Simulation Economic Modes

Responses to the choices within the DCE were analysed using a clustered conditional logit regression model (MNL) (McFadden 1974), within STATA 11.0. Willingness to pay was estimated for marginal changes in attributes and 95% confidence intervals obtained from bootstrapping with 1000 replications (Hole, 2007b). Observed preference heterogeneity was explored with respect to age.

The WTP values generated for the full group were then incorporated into a discrete event simulation (DES) model (Caro et al, 2010) The DES approach was chosen because of its flexibility in following individuals through monitoring, allowing for the chronological sequence of events, with an individuals' situation being permitted to change as a result of an event happening. For example, in the model presented here, this event may be a visit to the ophthalmologist to undergo monitoring tests for glaucoma, with the change occurring if a diagnosis of glaucoma occurred. Within the current DES, on entry, each individual had a predefined risk of developing open angle glaucoma (OAG). This was based upon

characteristics sampled from distributions for the population of interest (e.g. age, measures of vision such as intraocular pressure and cup to disk ratio, etc) and a risk prediction algorithm (Burr, Botello-Pinzon *et al*. 2012).

The model compared the consequences for each individual for five alternative surveillance pathways. Two surveillance pathways reflected current UK practice for monitoring ocular hypertension according to guidance developed by the National Institute of Health and Clinical Excellence (NICE); two pathways were developed from the outputs of the surveillance for ocular hypertension (SOH) study including the findings from the qualitative stage of the DCE described in this paper, (Burr, Botello-Pinzon et al. 2012), and the final treat all pathway reflected emerging findings from the literature (van Gestel et al. 2009). Thus, the surveillance pathways compared were: ***NICE intensive*** (four monthly to annual monitoring) and ***NICE conservative*** (six-monthly to biennially) with treatment according to baseline risk stratification on age, IOP and central corneal thickness (CCT); ***SOH hospital*** and ***SOH community***: monitoring biennially and treatment initiated for a ≥6% 5-year glaucoma risk, included repeated IOP measurements, within four months, following treatment initiation or change; ***Treat all***: treatment if intraocular pressure (IOP)>21mmHg, IOP measured annually (community optometry) and referral to hospital if inadequate treatment response.

Results

814 individuals responded to the questionnaire. All individuals were adults, 49% were males and half of the sample was over 45 years old. The sample was roughly evenly distributed in terms of education between those completing secondary, college and university level. Participants from England comprised 79% of the sample (645), 12% from Scotland (97), 6% from Wales (48), and 3% from Northern Ireland (24).

Number of choices

Over 68% passed Sen's contraction questions, while a further 24%, 6% and 1.4% failed one, two or three questions, respectively. Only three people failed 4 consistency check questions. From Table 3, fewer 'fails' occurred in the latter part of the questionnaire (5%) compared to 'fails' when the rationality question was at the start of the questionnaire (14%). This supports the idea of learning as respondents complete the DCE.

Table 3: Number of fails according to question number

Question Number	Number of fails	%
5	117	14%
15	92	11%
26	86	11%
34	43	5%

Regression results

All respondents were included in the regression analysis; results are available from the authors on request. Table 4 shows willingness to pay (WTP) results for the MNL model for full sample and age based subgroup analyses. For the full group, the positive WTP for the alternative specific constant (£28) implies a general preference for being monitored suggesting, everything else being equal, respondents would prefer to have a monitoring programme. Respondents were willing to pay £0.03 per year for a 1 in 10,000 person reduction in the number of people converting to glaucoma over the next 10 years. Thus, if 860 OHT persons per 10,000 avoided progressing to glaucoma within 10 years (as might occur when moving from no monitoring to a monitoring service), the WTP value would be £25.80 (£0.03 * 860) per year. As expected, unwanted effects of treatment reduce the value of any monitoring service. For example, WTP would be reduced by £16 if respondents experience some unwanted effects of treatment; £34 if they experience quite a lot of unwanted effects of treatment; and £59 if they experience severe unwanted effects of treatment. If individuals feel at ease and understood they would value the monitoring programme at an additional £46; experiencing one of these components would increase value by approximately £26.

The WTP age subgroup analysis provided better fitted models, suggesting age heterogeneity in preferences (results will be available at the conference). Those aged 50 and over have a stronger preference to be monitored, with the 50-65 age group having the strongest preference. These older age groups also had the strongest preference against unwanted side effects. However, they have the lowest preference for becoming visually impaired, with the preference being lower in the 50-65 age group compared to those aged 65 and over. This may reflect an increasing fear of ill health, and therefore blindness, as you get older (though it does not reach the levels of the younger group, who have the strongest

aversion to visual impairment, as expected). The 50 -65year age groups also valued more highly the provision of 'communication and understanding', with preferences being consistently higher than those who are 65 and older.

Table 4 Willingness to pay for subgroup analysis

	Willingness to pay (£) (95%CI)			
	Full model	**Age**		
Variable		**Under 50 years old**	**Between 50 and 65 years old**	**Over 65 years old**
Number of observations =	78144	44160	18624	15360
Clusters (individuals) =	814	460	194	160
Alternative specific constant (ref: no monitoring)	28 (19; 35)	23 (12; 32)	35 (17; 51)	38 (10; 60)
10-year risk of:				
conversion to glaucoma of individuals with ocular hypertension	-0.03 (-0.03; -0.02)	-0.03 (-0.03; -0.02)	-0.03 (-0.04; -0.02)	-0.03 (-0.05; -0.02)
progressing to glaucoma severe for OHT individuals	-0.06 (-0.09; -0.03)	-0.07 (-0.1; -0.03)	-0.02 (-0.08; 0.04)	-0.09 (-0.17; -0.02)
becoming visually impaired for OHT individuals	-0.65 (-0.9; -0.44)	-0.88 (-1.2; -0.61)	-0.17 (-0.62; 0.22)	-0.47 (-1.24; 0.13)
Unwanted treatment effects (ref: none):				
some	-16 (-19; -13)	-13 (-16; -10)	-16 (-24; -11)	-24 (-36; -16)
quite a lot	-34 (-39; -29)	-31 (-37; -25)	-36 (-48; -26)	-41 (-61; -27)
severe	-59 (-68; -51)	-53 (-63; -44)	-66 (-87; -50)	-74 (-109; -51)
Communication and understanding (ref: not feel at ease and did not understand)				
felt at ease and understood	47 (40; 54)	36 (30; 45)	62 (47; 82)	64 (45; 93)
felt at ease but did not understand	26 (22; 30)	22 (18; 27)	30 (22; 41)	35 (24; 51)
did not feel at ease but understood	26 (22; 31)	18 (14; 24)	37 (27; 50)	40 (27; 61)
Hospital setting (ref: community)	1 (-1; 3)	1 (-1; 3)	-1 (-6; 4)	5 (-1; 11)
Likelihood-ratio test:				
LR chi2(24) =	1846.23			
Prob > chi2 =	0.001			

Table 5 shows the CBA results for the full group. The SOH pathways provided on average £600 more benefit than the NICE based pathways. This can be explained by the added value generated by better communication and understanding assumed within the SOH pathways. Moreover, net-benefits for all the pathways other than the SOH hospital were

negative, indicating that only SOH hospital pathway would be considered worthwhile compared with the reference case of no monitoring. The main driver of these results is the very low valuation put on prevention of OAG and subsequent progression coupled with the low rates of OAG observed. Finally, it should be noted that the intervention with the minimum level of care - Treat all - had a negative net-benefit. This result is due mainly to the effect of the unwanted effects of treatment compared with the reference case where a proportion of these would be attended. In the Treat all programme everyone would initially receive medical treatment and hence the proportion of individuals suffering the unwanted effects of treatment is correspondingly greater than for the reference case.

Table 5 Cost-benefit analysis results (discounted)

Pathways	Average Total Cost (£)	Average Total Benefits (£)	Average Net Benefit (£)
Treat all	679	564	-114
SOH hospital	1241	1693	452
SOH primary care	1981	1484	-498
NICE conservative	2372	1051	-1321
NICE intensive	4148	1099	-3049

Discussion and conclusion

A DCE was conducted to elicit preferences for alternative glaucoma monitoring programmes for individuals with OHT. Qualitative work was used to inform attributes and levels and to develop the questionnaire, pictures were used to explain glaucoma and risk, a Bayesian design was employed, and the internet was used to collect data. Two points emerged which the authors would value discussion around:

How many choices to include in a DCE?

- We included 4 tests of rationality to provide some insight into the number of choices an individual can manage. Notably, fewer respondents failed the rationality tests as they proceeded through the questionnaire. This finding supports the idea of a learning effect but also challenges the notion of a fatigue effect stepping in after 16 choice

questions (de Bekker-Grob, Ryan et al. 2012). This does raise the question of whether health economists are being too cautious in our DCEs with respect to the number of choices. There is also the possibility that respondents are learning how to complete the task, as opposed to learning about their preferences. This is clearly an important area for future research.

Challenges when integrating the DCE and economic evaluation model **-** A number of issues had to be addressed to marry the DCE and DES model.

- Crucially, the DES model has to estimate, for a simulated sample of individuals, the time they would spend in all the states generated by the DCE. For example, the time that simulated individuals spent with (severe) unwanted effects of treatment or the number of individuals converting to glaucoma. However, mismatches between DCE attributes/levels and the elements incorporated into the DES model can occur. For example, in the hernia modelling (Vale, 2005) one important event, which was an important determinant in QALY scores, was the incidence of persisting numbness. An attribute for persisting numbness was not included in the DCE (because at the time the DCE was conducted it was not expected that there would be any differences in this outcome although subsequent evidence emerged that risk of persisting numbness did differ between treatments). Within the current application the DCE and DES were designed simultaneously to try to ensure compatibility. However, as the research progressed the design of the DES had to be refined to reflect the availability of data with which to model some parts of the process. For example, when including unwanted side effects in the DES we hoped to include some measure of intensity of side effects to mirror this attributes levels in the DCE. However, the data obtained from the literature on the proportion of individuals suffering unwanted effects of treatment did not separate between '*some*' and '*quite a lot*' of unwanted treatment effects. Therefore, two levels within the unwanted treatment effects (i.e. 'some' and 'quite a lot') were collapsed into a single level when incorporating this attribute into the DES model. The compatibility of the DCE and economic evaluation model is clearly important, and should be closely considered when incorporating DCEs into economic evaluation models.

- Another important issue is the time horizon chosen for the evaluation. In economic evaluations the modelling of life time costs and benefits is typically recommended because important differences between treatments may occur or persist over the entire lifetime of individuals. In the DCE conducted here a 10 year time horizon was used to allow for a better understanding of the risk figures. However, this meant important aspects of benefit may have been excluded from the DCE valuations. In particular, because glaucoma is a slowly progressing condition, very few people would progress from no glaucoma to visual impairment in 10 years. It is unclear whether respondents could have sensibly interpreted risks over a longer time horizon but what is clear is that more careful consideration of the adoption of any form of time frame in a DCE is needed to ensure its compatibility with the economic evaluation modelling.

- Given there is no requirement for individuals to take up surveillance an opt-out option was included in the DCE. This opt-out assumed no side effects of treatment. However, glaucoma may, under an opt-out option, be opportunistically detected and eventually treated and a proportion of those treated would then go on to experience treatment side effects. Ideally a more considered opt out would have been more fully defined.

- One of the key components of any economic evaluation is to explore under what conditions conclusions might change. In particular analysis is performed to explore whether conclusions are sensitive to the heterogeneity within the population studied. The econometric model used to analyses the DCE in this paper, the MNL model, may underestimate heterogeneity due to unobservable individuals' characteristics. To this end alternative econometric models have been suggested for analysing DCE data (e.g. mixed logit (MXL) (Hole 2007a, Eberth, Watson et al. 2009), latent class models (LCM) (Hole 2007a), and generalised MNL (Hole 2007a, Fiebig, Keane et al. 2009)). An advantage of MXL is that there is no requirement to define, a priori, the source of unobserved heterogeneity. However, the researcher has to decide which model parameters (attributes) are randomly distributed, and what distribution to impose. The source of unobserved heterogeneity is also not observed, so the output may be less useful to policy makers and for use within an economic evaluation (where the source of heterogeneity is required). LCM allows the researcher to define a number of classes

(groups) with alternative models fitted to each class, and probability of membership (according to characteristics of individuals) can be estimated. Whilst the use of different econometric approaches has received some attention in the health economics literature, consideration of whether more elegant econometric models have any practical value within an economic evaluation framework is an important area for discussion.

In conclusion, this paper adopted up to date DCE methods (qualitative work to inform attributes and levels and to develop the questionnaire, pictures to explain glaucoma and risk, a Bayesian design, and the internet to collect data) to look at preferences for alternative monitoring systems for individuals with ocular hypertension. The paper contributes at the methodological level at 2 levels. Firstly, through inclusion of rationality tests consideration was given to the optimal number of choices to include in a DCE. An interesting finding was that as individuals progressed through the DCE they failed fewer rationality tests. Whilst further research is needed in this area our findings do suggest that Health Economists have been too cautious, often restricting the number of choices to 16 or less. The main contribution was to discuss issues raised in the incorporation of DCE results into an economic evaluation modelling framework. More specifically, in this paper, the results of the DCE were incorporated into a DES economic evaluation model. The importance of marrying the needs of the two approaches (DCE and DES) at the outset, time horizons employed; clearly defining the opt-out option (common to all DCEs) and issues around the appropriate econometric analysis were all discussed. Both issues are clearly important in terms of future research.

Acknowledgements

This work was part of a project commissioned by the NIHR HTA Programme, UK (Project number 07/46/02). We would like to thank the surveillance for ocular hypertension study group (Health Technol Assess 2012: forthcoming), Barbara Eberth & Verity Watson for advice on econometric analysis and discrete choice experiments, Ann Hoste on behalf of the Belgian Glaucoma Society for permissions to use their pictorial representation of glaucoma, and the patients who participated in the focus group and the Advisory panel for their invaluable contribution. The Health Services Research Unit, and Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen are both core-funded by the Chief Scientist Office of the Scottish Government Health Directorates.

References

ANCKER, J.S., SENATHIRAJAH, Y., KUKAFKA, R. and STARREN, J.B., 2006. Design features of graphs in health risk communication: a systematic review. *Journal of the American Medical Informatics Association*, 13(6), pp. 608-618.

BECH, M., KJAER, T. and LAURIDSEN, J.+., 2011. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Economics*, 20(3), pp. 273-286.

BLIEMER, M.C.J., ROSE, J.M. and HENSHER, D.A., 2009. Efficient stated choice experiments for estimating nested logit models. *Transportation Research Part B: Methodological*, 43(1), pp. 19-35.

BURR, J.M., KILONZO, M., VALE, L. and RYAN, M., 2007. Developing a Preference-Based Glaucoma Utility Index Using a Discrete Choice Experiment. *Optom Vis Sci*, 84(8), pp. 797-808.

BURR, J., BOTELLO-PINZON, P., TAKWOINGI, Y., HERNÁNDEZ, R., VAZQUEZ-MONTEZ, M., ELDERS, A. and ET AL., 2012. Surveillance for ocular hypertension: evidence synthesis and economic evaluation. *Health Technol Assess*, 16. in press

CARO JJ, MÖLLER J, GETSIOS D. 2010 Discrete event simulation: the preferred technique for health economic evaluations? Value Health. 2010 Dec;13(8): 1056-60

CAUSADE, S., ORTUZA, J.D.D., RIZZI, L. and HENSHER, D., 2005. Assessing the influence of design dimensions on stated choice experiments estimates. *Transportation Research*, Part B, pp. 621-640.

COAST, J., 2006. Maximising responses to discrete choice experiments: a randomized trial. *Applied Health Economics & Health Policy*, 5(4), pp. 249-260.

DE BEKKER-GROB, E.W., RYAN, M. and GERARD, K., 2012. Discrete choice experiments in health economics: a review of the literature [epub ahead of print]. *Health Economics*, 21, pp. 145-172.

EBERTH, B., WATSON, V., RYAN, M., HUGHES, J. and BARNETT, G., 2009. Does One Size Fit All?: Investigating Heterogeneity in Men's Preferences for Benign Prostatic Hyperplasia Treatment Using Mixed Logit Analysis. *Medical Decision Making*, 29(6), pp. 707-715.

EDWARDS, A. and ELWYN, G., 2001. Understanding risk and lessons for clinical risk communication about treatment preferences. *Quality in Health Care*, 10(Suppl I), pp. i9-i13.

FAGERLIN, A., ZIKMUND-FISHER, B.J. and UBEL, P.A., 2011. Helping patients decide: ten steps to better risk communication *JNCI J Natl Cancer Inst*, 103(19), pp. 1436-1443.

FIEBIG, D.G., KEANE, M.P., LOUVIERE, J. and WASI, N., 2009. The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science*, .

HANLEY, N., WRIGHT, R. and KOOP, G., 2002. Modelling recreational demand using choice experiments: rock-climbing in Scotland. ***Environmental and Resource Economics*****, 22(3), pp. 449-466.**

HENSHER, N., STOPHER, P. and LOUVIERE, J., 2001. An explanatory analysis of the effect of number of choice sets in designing choice experiments: an airline choice application. ***Journal of Air Transoport Management*****, 7, pp. 373-379.**

HERNANDEZ, R.A., BURR, J.M. and VALE, L.D., 2008. Economic evaluation of screening for open-angle glaucoma. ***International Journal of Technology Assessment in Health Care*****, 24(2), pp. 203-211.**

HOLE, A.R., 2007a. ***Modelling Heterogeneity in Patients' Preferences for the Attributes of a General Practitioner Appointment.*** **SSRN.**

HOLE, A.R., 2007b. A comparison of approaches to estimating confidence intervals for willingness to pay measures. ***Health Economics*****, 16(8), pp. 827-840.**

KUHFELD, W.F., 2010-last update, Experimental design: efficiency, coding, and choice design [document on the Internet] [Homepage of February 2011], [Online].

KUHFELD, W.F., TOBIAS, R.D. and GARRATT, M., 2010-last update, Efficient Experimental Design with Marketing Research Applications [document on the Internet] [Homepage of February 2011], [Online].

LLOYD, A.J., 2001. The extent of patients' understanding of the risk of treatments. ***Quality in Health Care*****, 10, pp. i14-i18.**

LOUVIERE, J., 2004. Random utility theory based stated preference elicitation methods: an application in health economics with special reference to combining sources of data. ***Centre for the Study of Choice (CenSoC)*****, Working Paper No. 04-001.**

MCFADDEN, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In: P. ZAREMBKA, ed, ***FRONTIERS IN ECONOMETRICS.*** **New York: Academic Press, pp. 105-142.**

PEARMAIN, D., ed, 1991. ***Stated preferences techniques: guide to practice.*** **The Hague: Steer Davis Gleave Group.**

REGIER, D., 2008. ***Discrete choice experiments informing cost benefit analysis: a Bayesian approach with an application to genetic testing. [Ph.D. thesis].*** **Aberdeen: University of Aberdeen.**

RITCHIE, J. and SPENCER, L., 1994. Qualitative data analysis for applied policy research. In: A. BRYMAN and R.G. BURGRESS, eds, ***Analysing Qualitative Data.*** **London: Routledge, pp. 173-194.**

ROSE, J.M. and BLIEMER, M.C.J., 2008. Stated preference experimental design strategies. In: D.A. HENSHER and K.J. BUTTON, eds, ***Handbook of Transport Modelling.*** **Oxford: Elsevier, pp. 151-180.**

ROSE, J.M., BLIEMER, M.C.J., HENSHER, D.A. and COLLINS, A.T., 2008. Designing efficient stated choice experiments in the presence of reference alternatives. ***Transportation Research Part B: Methodological*****, 42(4), pp. 395-406.**

SEN, A., 1993. Internal Consistency of Choice. ***Econometrica*****, 61(3), pp. 495-521.**

STOPHER, P. and HENSHER, D., 2000. Are more profiles better than fewer? Searching for parsimony and relevance in stated choice experiments. ***Research Transoportation Record 1719*****, Paper No 00-1255.**

VALE, L.D., 2005. ***The progressive economic evaluation of a surgical technique: a case study of surgical repair of inguinal hernia. [Ph.D. thesis].*** **Aberdeen: University of Aberdeen.**

VAN GESTEL A, SEVERENS H, WEBERS C, BECKERS H, JANSONIUS N, SCHOUTEN J. A novel approach to establish at which IOP to start treatment for ocular hypertension. World Glaucoma Congress, Boston, 2009.