CHERE
CENTRE FOR HEALTH ECONOMICS
RESEARCH AND EVALUATION

UNIVERSITY OF
TECHNOLOGY SYDNEY

UTS:CHERE

The Centre for Health Economics Research and Evaluation (CHERE) was established in 1991. CHERE is a centre of excellence in health economics and health services research. It is a joint Centre of the Faculties of Business and Nursing, Midwifery and Health at the University of Technology, Sydney, in collaboration with Central Sydney Area Health Service. It was established as a UTS Centre in February, 2002. The Centre aims to contribute to the development and application of health economics and health services research through research, teaching and policy support. CHERE's research program encompasses both the theory and application of health economics. The main theoretical research theme pursues valuing benefits, including understanding what individuals value from health and health care, how such values should be measured, and exploring the social values attached to these benefits. The applied research focuses on economic and the appraisal of new programs or new ways of delivering and/or funding services. CHERE's teaching includes introducing clinicians, health services managers, public health professionals and others to health economic principles. Training programs aim to develop practical skills in health economics and health services research. Policy support is provided at all levels of the health care system by undertaking commissioned projects, through the provision of formal and informal advice as well as participation in working parties and committees.

University of Technology, Sydney
City campus, Haymarket
PO Box 123 Broadway NSW 2007
Tel: +61 2 9514 4720
Fax: + 61 2 9514 4730
Email: mail@chere.uts.edu.au
www.chere.uts.edu.au

Equity weights for economic evaluation: An Australian Discrete Choice Experiment

Richard Norman*[1] and Gisselle Gallego[1]

CHERE WORKING PAPER 2008/5

* Corresponding author:

1. Centre for Health Economics Research and Evaluation
Faculty of Business
University of Technology, Sydney
PO Box 123, Broadway, Sydney, 2007 Australia

First Version: September 2008
Current Version: September 2008

CHERE
CENTRE FOR HEALTH ECONOMICS
RESEARCH AND EVALUATION

UNIVERSITY OF
TECHNOLOGY SYDNEY

Abstract

Background: Economic evaluation of healthcare interventions and technologies using the quality-adjusted life year (or life year) usually values outcomes independently of who they accrue to. This is a simplifying assumption relating to the more complex societal preferences. While the premise of equal value has been criticised as being unreflective of societal views, no alternative has gained significant traction.

Aims: To identify the trade-offs made by an Australian population between total gain in life expectancy, initial life expectancy, gender, income and smoking status, and then to generate equity weights for economic evaluation from these results.

Method: A discrete choice experiment was used in an online panel. 241 respondents answered twelve binary choices, and the results were analysed using logistic regression. Equity weights were then generated using Hicksian compensating variation.

Results: A typical individual was willing to discriminate based on smoking and income, but not on gender or initial life expectancy (although the last of these is considered within a narrow range of 55-75 years). However, there was considerable heterogeneity in respondents. Equity weights ranged from 0.673 for smokers with an above average income to 1.207 for non-smokers with a below average income. This result was sensitive to the point at which the marginal utility of time was estimated.

Conclusion: Healthcare decision making, using an orthodox QALY model, does not capture the views of society, particularly with regard to smoking or income. We have presented an alternative approach, weighting outcomes dependent on the personal characteristics of the individual receiving them. The feasibility of including this finding in economic evaluation is as yet uncertain and has to be investigated further.

Introduction

In his 1997 paper, Alan Williams stated that "(D)*eath at 25 is viewed very differently from death at 85*." (Williams, 1997). This idea is fairly universal across people and cultures, and is perhaps based on the Aristotelian view that what is important is to allow all an opportunity to flourish. While the 85 year old will probably desire to continue living, they are much more likely to have had the opportunity to pursue their goals and it is arguable that a society should reflect the relative importance of the 25-year old being given that same opportunity in our decisions and structures. In contemporary decision-making in health and healthcare, economic evaluation is considered alongside issues of equity (and others) and the balance between these often conflicting issues is unclear, arguable necessarily so.

The current orthodoxy in economic evaluation of healthcare is grounded in utilitarianism (although Williams argues the use of QALYs does not necessitate this (Williams, 1996), only differing in that is seeks to maximise health, rather than utility associated with health. However, since comparison between people with regard to the utility derived from health is difficult, this simplifying assumption is probably necessary.

Conventionally, the cost-effectiveness of an intervention relative to another is defined by the incremental cost-effectiveness ratio (ICER), which can be defined over a population *P*. Within *P* each individual *i* expects a health gain g_i from an intervention. The cost of providing the intervention in that individual is x_I, and is often constant across individuals. Conventional economic evaluation would evaluate the incremental cost-effectiveness ratio as

$$ICER = \frac{\sum_{i=1}^{P} x_i}{\sum_{i=1}^{P} g_i}$$

This can be broken down into analysis for various subpopulations, potentially differing in terms of their expected costs and expected outcomes. The application of utilitarianism captured by the conventional use of QALYs has been criticised on the grounds it has the potential to be ageist, sexist or racist (Sassi, et al., 2001). There has been significant criticism of the QALY model as representative of population values (Bryan, et al., 2002, Roberts, et al., 1999). However, alternatives to utilitarianism have so far failed to gain significant traction. Rawls applied the 'Maximin principle' to primary social goods. As Sassi notes, while Rawls did not include health in this category, it is a feasible perspective to take, asserting that the aim of the health sector should be to improve the health of the person with the worst health (Sassi, et al., 2001). However, this is likely to be an extreme view. While utilitarianism treats gains to different people as perfect substitutes, Maximin treats them as perfect complements.

While likely to prioritise towards those with the worst expected health, an accurate representation of a social welfare function is likely to also value health gains accrued to others. Using the previous notation, the ICER for the intervention across all groups under a system of equity weighting will be

$$ICER = \frac{\sum_{i=1}^{P} x_i}{\sum_{i=1}^{P} w_i g_i}$$

where w_i is the weight associated with each individual *i* and is a function of a number of individual-level characteristics, which might include gender, income, smoking status, life expectancy, parenthood (or having other dependents) and many other potentially relevant areas. As before, this can be broken into sub-populations, identifying not just the different capacity to, and likelihood of benefit, but also the value that society places on the individual level gain. It is useful to set the mean of w_i across the entire population at 1, so an intervention considered over an entire population will have an equal ICER both with and without equity weighting. The difference emerges when an intervention affects one of more of the population sub-groups differently, or is likely to be targeted at a population group with levels of (for example) income or smoking different from the general population. It is the primary aim of this paper to present one way of using stated preference experiments to generate values for w_i.

Existing evidence in the area has suggested that w_i may differ from 1 for different individuals (i.e. society is willing to discriminate between individuals). A paper from the United Kingdom has investigated how comparable NHS clinicians were to members of the public with regard to how they value equality (Tsuchiya and Dolan, 2007). They posed a series of questions, one option increased the life expectancy of both the highest and lowest social classes by the same amount, and one which targeted the gains towards those in the lowest social class. Both populations they considered broke down into three sections, those who prefer whichever option maximises total gain, those who prefer targeting irrespective of what gains are offered, and those who follow either behaviour dependent on the scale of the gains. The general population sample suggested approximately half of the sample prioritised solely on the basis of gains in life expectancy (i.e. in keeping with the conventional use of QALYs) while 50% did not.

In this study, we have chosen to focus on equality of life expectancy. People's preferences with regards to equality are likely to be complicated, including not just equality of outcome but also equality in other areas, such as gain and access. The task of modelling these attitudes is important and interesting, but beyond the scope of this work. However, it is noteworthy that work from the United Kingdom has identified that most egalitarians are egalitarians in outcome space rather than in gains space (with the possible exception of gains between genders) (Tsuchiya and Dolan, 2008). One issue with the use of life expectancy rather than quality-adjusted life expectancy is that quality of life assumptions may leak into life expectancy (e.g. people may discriminate against smokers because they assume the quality of life of extra years will be relatively poor).

Discrete choice analysis, which is outlined more thoroughly below, is a method for investigating these kinds of complex thought processes, and allows flexible functional forms to be evaluated. The use of conjoint analysis to set priorities in an Australian population has been established (Browning and Thomas, 2001, Jan, et al., 2000), as has the desire of the Australian population to be involved in decision-making (Wiseman, et al., 2003). The aims of the project are: To identify how important characteristics of both the society (in terms of initial inequality of life expectancy) and of individuals affect the trade-offs made; and to show how these views might be integrated into economic evaluation of health and healthcare.

Method

The Discrete Choice Experiment (DCE) was presented as a choice between two unlabelled healthcare programs, each of which benefited a different group of hypothetical individuals and increases the life expectancy in the target group by between 2 and 14 years. The groups were composed of people aged 40, and defined in terms of their initial life expectancy, their gender, smoking status and income. Respondents to the survey were identified through an online panel of Australian residents, and the experiment was run through an online survey company (Survey Engine Pty). The variables for the construction of the set of w_i are shown in Table 1. We do not claim that the population sub-divisions investigated here are exhaustive: It may be that other individual level characteristics might affect the value society places on an individual's increase in life expectancy. For example, Olsen and colleagues identify a range of other criteria over which people have been shown to be willing to discriminate (Olsen, et al., 2003). However, it is essential that stated preference experiments balance the exhaustiveness of the issues investigated with respondent fatigue (and subsequent inaccuracies in responses).

Table 1: Variables in the Analysis

Variable	Possible levels	Details	Coding
Initial life expectancy	5	55-75 years in intervals of 5	Continuous
Gain in life expectancy	7	2-14 years in intervals of 2	Continuous
Income	2 for each group	Average / below average	Dummy
Smoking status	2 for each group	Yes / No	Dummy
Gender	2 for each group		Dummy

The design for the DCE was generated using software publicly available from Burgess (Burgess, 2007). With the parameters outlined, it was possible to produce a fractional design of 8 blocks of 12 choice sets which estimate the main effects and all two factor interactions independently (i.e. no correlation between estimates). As two factor interactions were designed for, it was not possible to identify the efficiency of the design as no clear method for doing so has been acknowledged as yet. Prior to running the experiment with the online panel, we undertook a pilot in a convenient sample of academic peers (n=20) and students (n=30) looking at the ease and comprehensibility of the task. This included a mock-up of a choice set as shown in Figure 1.

Figure 1: A Sample Choice Set

Look at the two possible healthcare programs below. They will both help 100 people. All of these people are 40 years of age. They currently have a particular life expectancy, and all people in each group have the same gender, the same income (either below or above average), and they all either are or are not current smokers. Please consider the groups involved and the gain in life expectancy. At the end of the scenario, you will be asked to prioritise one of the two programs.

	Program A	Program B
Current age (in years)	40	40
Life Expectancy without program (in years)	70	60
Gain in Life Expectancy Under Program	12 years	2 years
Gender	Male	Male
Income	Below average	Above average
Smoking status	No	Yes
Which of these two programs would you choose to prioritise?		

prev next

The information provided by the pilot groups were integrated into the survey, which was then run online. The use of online panels is a potentially valuable addition to survey methodology as the possible sample size associated with a fixed budget is considerably higher. Initial investigations have identified equivalence in response between online and non-online respondents (Gwaltney, et al., 2008) although it remains important to compare the co-variates in the two population groups to minimise divergence due to difference in populations. Comparisons between the study group and the general population were made using χ^2 goodness of fit tests. We also asked how easy or difficult the respondent found the task using a scale of 1-5 (with 1 being very difficult, and 5 being very easy), and used regression (OLS when the scale was continuous, logistic when the scale was transformed to categorical data) to identify whether different types of people found the task particularly difficult.

Regarding the analysis of responses, as there are only two choices for the respondent, binary choice models can be employed. These are based (as are more advanced specifications such as mixed logit) on the random utility model described by McFadden (McFadden, 1981). The utility for individual *i* of alternative *j* in scenario *s* is defined in the following way:

$$U_{isj} = X_{isj}^{'}\beta_i + \varepsilon_{isj}$$

where β_i is a vector of co-efficients and X_{isj} is a vector of explanatory variables. If we assume the error term to be identically and independently distributed as extreme value, we get the standard general multinomial logit (MNL) specification in which the probability that the individual chooses alternative *j* in scenario *s* is defined as:

$$P_{isj} = \frac{\exp(X_{isj}^{'}\beta_i)}{\sum_h \exp(X_{ish}^{'}\beta_i)}$$

As there are only two choices in the binary choice model , we can define *k* as the alternative which is not *j* and then simplify this to:

$$P_{isj} = \frac{\exp(X_{isj}^{'}\beta_i)}{\exp(X_{isj}^{'}\beta_i) + \exp(X_{isk}^{'}\beta_i)}$$

To estimate the marginal effect of each of the attributes, we adopted the Hicksian compensating variation approach of Small and Rosen (Small and Rosen, 1981), which was advocated more recently in healthcare by Lancsar and Savage (Lancsar and Savage, 2004). The expression for the compensating variation, accounting for uncertainty regarding which alternative will be selected in a DCE is

$$CV = \frac{1}{\lambda}\left[\ln \sum_{j=1}^{J} e^{V_j^0} - \ln \sum_{j=1}^{J} e^{V_j^1}\right]$$

where λ is the marginal utility of additional life expectancy, v_j^0 and v_j^1 are the values of the indirect utility function for each j before and after the quality change (e.g. with or without a particular characteristic), and *J* is the number of options in the choice set. In this investigation, it was necessary to test whether λ is linear as this QALY model assumption has been criticised as unrealistic (Dolan, et al., 2005, Viney and Savage, 2006). If there is a statistically significant quadratic term (representing increasing or (more likely) decreasing marginal utility of extra years of life, it will be necessary to estimate λ at a point (as the slope of a quadratic changes over the independent variable). If necessary, we will use an arbitrary value of 10 years, and identify how sensitive the consequent weights are to this assumption.

In this context, the Compensating Variation approach provides an estimate of the number of years gain (for example) for a smoker is equivalent to the number of years gain for a non-smoker. To construct equity weights from this, we then imply a time trade-off type decision. If an additional 10 years of life for a group of smokers is valued equally to an additional eight years of life for a non-smoker group, the equity weight for smokers relative to non-smokers would be 8/10 = 0.8.

It should be noted that there has been some disagreement about the appropriateness of the Small and Rosen approach in healthcare (Ryan, 2004, Santos Silva, 2004). Additionally, there are a variety of alternative approaches to the estimation of relative attribute impact that have been considered in the literature (Lancsar, et al., 2007). Therefore, we present the results in such a way as to allow estimation of the trade-offs in alternative ways. We present equity weights for each population group (as smoking, income and gender are assumed to be binary, a maximum of 8 weights will be presented). The reason for separating them (rather than presenting a weight for smokers, a weight for low-income groups etc) is that it is possible interactions terms will prove significant, thus making the latter approach impossible. Only preferences for targeting that are statistically significant are included.

As the betas are estimated with a particular population acting as baseline (in this case, male, non-smokers with above average incomes), it is helpful to correct the equity weights derived so the average population weight is 1. The reason for this is that it allows economic evaluation with and without equity weights to be judged using the same particular thresholds (for example the often quoted £30,000 per QALY in the United Kingdom) (Devlin and Parkin, 2004, National Institute for Health and Clinical Excellence, 2008). This correction will be done by identifying the proportions of the sample in each of the eight possible population groups (for example, female, non-smokers with a below average income would be one) in our sample. For these purposes, it is assumed that the cut-off between average and below average gross household incomes is $60,000 per annum. This adjustment will retain a common ratio between weights for population groups.

Finally, demographics were introduced into the model. As the analysis is a conditional logistic regression, this required setting a base individual (in this case male, younger than 40, household income below $50,000 per annum, non-smoker) and estimating the interactions between divergences from this base individual and the various attributes. Following this, sub-group analysis was undertaken to illustrate whether different groups of individuals discriminate differently. As the respondent information was collected to match the categories given in the experiment (so smoking status, income and gender), the analysis was repeated looking at differences in emphasis between groups, particularly in respect to their attitude towards their defining characteristic. For example, smokers and non-smokers are analyses separately to investigate whether one has a different attitude to discrimination for or against smoking.

<u>Results</u>

Respondents

324 respondents began the study, of which 241 completed (giving a completion rate of 74.4%). The median time spent answering the survey was 383 seconds (inter-quartile range of 284-500). Details of the respondents, alongside Australian population norms, are given in Table 2. The sample has a larger proportion of males, mid-aged people, and people with higher levels of education than the general population.

Table 2: Demographic Characteristics of Respondents

Characteristic	Value / Range	Sample	Population[3]	P-value*
Gender[1]	Female	47.50%	56.09%	0.0084
Age (years)	16-29	11.84%	21.33%	
	30-44	21.17%	23.98%	
	45-59	30.41%	22.40%	<0.0001
	60-74	32.09%	14.00%	
	75+	4.49%	18.29%	
Smoker	Yes	19.92%	23.00%	0.2329
Highest level of education	Primary	3.66%	40.51%	
	Secondary	33.80%	20.00%	<0.0001
	Trade certificate	37.80%	22.24%	
	Bachelor's degree or above	24.73%	17.26%	
Gross household income[2]	<$20,000	11.91%	15.77%	
	$20,000 - $40,000	20.70%	23.02%	0.4047
	$40,001 - $60,000	21.43%	17.64%	
	$60,001 - $80,000	13.98%	13.87%	
	$80,001 - $100,000	13.02%	11.03%	
	$100,001 +	18.94%	18.67%	

* Ho: Distribution identical in sample and population
[1] One individual had missing data for this variable
[2] 38 individuals chose to not disclose income
3 All data sourced from ABS (Australian Bureau of Statistics, 2002, Australian Bureau of Statistics, 2005, Australian Bureau of Statistics, 2006, Australian Bureau of Statistics, 2007)

Method of response

Tsuchiya and Dolan noted (Tsuchiya and Dolan, 2007) that proportions of the general population prioritised between alternative healthcare programs solely on efficiency grounds (42.4%) or on equality (9.2%). As the experiment undertaken here included more variables (such as smoking status) which might prevent domination by one domain (i.e. gain in life expectancy or initial life expectancy), we would not expect to replicate these figures. For comparison, the figures in our general population Australian survey are 24.1% for efficiency, and 4.4% for equality.

32 of the 241 (13.3%) respondents found the task difficult or very difficult. Dummy coding each of the levels of income, education, age, gender and smoking given in Table 2, OLS regression suggested no relationship (at the 5% level of significance) between type of person and ease of answering. When a categorical variable was created identifying those who found the task difficult or very difficult, logistic regression again identified no relationship between respondent characteristic and ease of response. This may be supportive evidence for the use of online surveys in the entire population (including those considered to be less computer literate), although it is arguable that only computer literate individuals would be a member of an online panel.

Betas from DCE experiments

The estimates of the parameters are provided in Table 3. Respondents preferred options with a greater gain in life expectancy, and chose non-smokers in preference to smokers. However, the initial life expectancy was not statistically significant, which is an unexpected result.

Table 3: Logistic Regression Results

	Model 1		**Model 2**		**Model 3**		**Model 4**	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Gain in life expectancy (years)	0.1955*	0.000	0.1969*	0.000	0.1971*	0.000	0.1288*	0.000
Gain in LE squared	-0.0043*	0.043	-0.0042*	0.047	-0.0043*	0.047		
Initial life expectancy (years)	0.0051	0.222	0.0045	0.294	0.0045	0.292	0.0045	0.293
Smoking			-0.5630*	0.000	-0.4257*	0.000	-0.4256*	0.000
Female			-0.0078	0.895	-0.0065	0.913	-0.0072	0.903
Below average income			0.1962*	0.007	0.3366*	0.000	0.3380*	0.000
Smoking x Below average income					-0.2798*	0.014	-0.2798*	0.014
Log-likelihood	-1838.5		-1788.1		-1785.1		-1787.04	
Pseudo R-squared	0.1278		0.1517		0.1531		0.1522	

* Significant at the 5% level
** Significant at the 10% level

Four functional forms were used, as outlined above. Model 1 tested the linearity of time, and suggested a diminishing marginal utility to time (which is at odds with the standard use of QALYs). It also represents symmetrical preferences, in that the only parameters are those that deal with inequality of life expectancy (therefore excluding individual characteristics). Model 2 introduced the characteristics of the hypothetical populations. It is noteworthy that the model fit improves by doing so, suggesting symmetrical preferences are unrealistic. Model 3 includes the only two-way interaction term which proved statistically significant (at the 5% level). Respondents were unwilling to discriminate on the basis of either initial life expectancy or gender. Respondents were willing to discriminate in favour of low income individuals, but this effect was largely negated if the potential gainers were smokers. Model 4 repeats Model 3, but excludes the statistically significant quadratic term. Doing so did not cause a large change in the log-likelihood (or Pseudo-R^2). Therefore, it might be dropped if it proves difficult for the estimation of equity weights.

Equity Weights for Economic Evaluation

The implicit equity weights using all 241 respondents are shown in Table 4.

Table 4: Equity Weights Derived From Statistically Significant Parameters

Smoking status	Income	Equity Weight relative to non-smoker, earning above the average	Population proportion (from the sample)	Corrected equity weight	Range
Smoker	Above average	0.700	8.37%	0.673	0.509-0.855
	Below average	0.739	11.99%	0.711	0.566-0.870
Non-smoker	Above average	1.000	42.07%	0.962	0.945-0.980
	Below average	1.255	37.57%	1.207	1.096-1.309

We repeated the analysis with the marginal utility of additional life expectancy judged at 2 and 14 years (i.e. the extreme values considered in our experiment). The equity weights under these two alternate methods differed, and are shown in brackets in the column headed 'corrected equity weights'. This is important for the use of these weights in economic evaluation as the point at which we appraise the marginal utility of additional life expectancy is arbitrary.

Sub-group analysis

The result of including demographics is given in Table 5.

Table 5: Introducing Demographics

Variable	Demographic (Income and Age excluded as NS)	Co-efficient	P-Value
Gain in LE*		0.2028	0.000
(Gain in LE)^{2*}		-0.0046	0.033
Initial LE		0.0051	0.235
Smoking*		-0.4797	0.019
Female*		-0.5331	0.000
Low Income		0.3120	0.165
Low Income x Smoker		-0.2592	0.371
Smoking*	Smoker	0.5669	0.005
Female		0.1268	0.376
Low Income		0.3097	0.182
Low Income x Smoker		0.1096	0.701
Smoking	Post-secondary qualifications	-0.1248	0.478
Female*		0.4495	0.000
Low Income		0.1448	0.465
Low Income x Smoker		0.0778	0.758
Smoking	Female	0.0554	0.743
Female*		0.5165	0.000
Low Income		-0.0598	0.755
Low Income x Smoker		0.0832	0.730
Log-likelihood	-1743.6	Pseudo-R^2	0.1659

* Significant at 5% level

Age and income as demographics were excluded as the interaction of each with any of the co-efficients were not statistically significant at the 5% level. The interactions which were significant at the 5% level were smoking status affecting the impact of smoking, gender affecting the impact of gender, and having post-secondary education affecting the impact of gender. The first two are likely to be situations in which individuals discriminate in favour of people similar to themselves, while the final statistically significant interaction is difficult to explain.

Sub-group analysis, presenting separate conditional logistic regression based on smoking status is given in Table 6.

Table 6: Logistic regression results by smoking status

	Smokers (n=48)			**Non-smokers (n=193)**		
	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value
Gain in life expectancy (years)	0.2521*	0.0848	<0.003	0.1817*	0.0393	<0.0001
Gain in LE squared	-0.0079	0.0051	0.121	-0.0031	0.0024	0.187
Initial life expectancy (years)	0.0210*	0.0093	0.024	0.0000	0.0049	0.984
Smoking	**-0.0004**	**0.1791**	**0.998**	**-0.5428***	**0.0918**	**<0.0001**
Female	-0.0192	0.1307	0.883	-0.0125	0.0665	0.851
Below average income	0.5277*	0.2087	0.011	0.2802*	0.1034	0.007
Smoking x Below average income	-0.2035	0.2531	0.421	-0.2955*	0.1286	0.022

* Significant at the 5% level
** Significant at the 10% level

The major difference between the smokers and the non-smokers is in bold. For smokers, smoking status has no impact on the choice between groups. For non-smokers, the z-score is -5.91 and significant at all commonly applied levels of significance. This result might be explained in two ways. On the one hand, self-interest may be influencing decisions, leading smokers to not discriminate against smokers (but not actively favouring them either) and non-smokers to discriminate heavily. Another possibility is that respondents are also considering the revealed attitude of smokers and non-smokers towards their own health. Arguably, those engaging in activities likely to damage health have a relatively lower valuation of that health. We can investigate which argument is more persuasive by considering an attribute (gender) which is not influenced by personal choice. This is shown in Table 7.

Table 7: Logistic regression results by gender

	Females (n=114)			**Males (n=126)**		
	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value
Gain in life expectancy (years)	0.1305*	0.0124	<0.0001	0.1261*	0.0121	<0.0001
Initial life expectancy (years)	0.0076	0.0061	0.211	0.0010	0.0061	0.872
Smoking	-0.3996*	0.1164	0.001	-0.4522*	0.1138	<0.0001
Female	**0.1698***	**0.0863**	**0.049**	**-0.1778***	**0.0814**	**0.029**
Below average income	0.2604**	0.1380	0.059	0.4050*	0.1256	0.001
Smoking x Below average income	-0.2493	0.1687	0.139	-0.3245*	0.1568	0.038

* Significant at the 5% level
** Significant at the 10% level

As with smoking, respondent gender influence choices. Both co-efficients are statistically significant, but in opposite directions meaning each gender relatively favours their own gender. In an area over which individuals have no choice, the discrimination against those that are different remains.

Discussion

We have begun to explore how an Australian population makes trade-offs between efficiency and equality of life expectancy, and have presented equity weights based on smoking status, income, gender, and life expectancy. Smoking status and income of potential recipients of healthcare are statistically significant predictors of choices between competing healthcare programs. Gender is not, nor is initial life expectancy. However, this last finding may be a consequence of the relatively narrow range of life expectancy considered here. It would be interesting to explore whether this non-statistically significant relationship remained when we ask people about gains in life expectancy for young adults or children. A further helpful addition to the experiment presented here would be to consider other factors which might influence societal valuation of gains to individuals, such as employment and whether the hypothetical person has dependents.

There are a number of strengths of the approach used here, and of the use of equity weights in economic evaluation. For the former, while respondent fatigue is an issue, discrete choice experiments offer a way of identifying multiple interacting effects which influence the choices people make. While the use of stated preference (rather than revealed preference) techniques is likely to be necessitated by the area under discussion, it should also be noted that we are assuming that stated preferences reflect actual preferences accurately. The use of equity weights is potentially helpful as it allows decision makers to reflect the views of the community they are serving. At present, the use of the QALY model in economic evaluation is a necessary simplification of societal views, but the inclusion of equity concerns into economic evaluation may allow a more complex and representative view to be adopted.

It should be noted that this use of discrete choice experiments to produce a representative societal valuation of outcomes appears to contradict Arrow's Theorem (Arrow, 1950). Arrow showed that it was not possible to aggregate individual welfare functions into a societal welfare function without violating one or more of a set of intuitive assumptions, namely non-dictatorship, universality, independence of irrelevant alternatives, and Pareto efficiency. In this paper, we have followed the direction taken in much of the literature (that is to maximise the predictive value of the model), but it should be noted that the DCE literature has not yet presented a clear rebuttal to Arrow's Theorem. Approaches in development, such as latent class modelling accounting for different error variances may offer a way forward.

References

Arrow KJ. 1950. A difficulty in the concept of social welfare. *Journal of Political Economy* **58**: 328-346.

Australian Bureau of Statistics. *Education and training indicators, Australia, 2002.* Canberra, 2002.

Australian Bureau of Statistics. 3222.0 - population projections, Australia, 2004 to 2101 2005.

Australian Bureau of Statistics. *Tobacco smoking in Australia: A snapshot, 2004-05* Australian Bureau of Statistics: Canberra, 2006.

Australian Bureau of Statistics. *Household income and income distribution, Australia, 2005-06.* Canberra, 2007.

Browning CJ, Thomas SA. 2001. Community values and preferences in transplantation organ allocation decisions. *Social Science and Medicine* **52**: 853-861.

Bryan S, Roberts T, Heginbotham C, McCallum A. 2002. QALY-maximisation and public preferences: Results from a general population survey. *Health Economics* **11**: 679-693.

Discrete choice experiments, Department of Mathematical Sciences, University of Technology, Sydney, available from http://crsu.Science.Uts.Edu.Au/choice [program], 2007.

Devlin N, Parkin D. 2004. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* **13**: 437-452.

Dolan P, Shaw R, Tsuchiya A, Williams A. 2005. QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics* **14**: 197-208.

Gwaltney CJ, Shields AL, Shiffman S. 2008. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. *Value in Health* **11**: 322-333.

Jan S, Mooney G, Ryan M, Bruggemann K, Alexander K. 2000. The use of conjoint analysis to elicit community preferences in public health research: A case study of hospital services in South Australia. *Australian and New Zealand Journal of Public Health* **24**: 64-70.

Lancsar E, Louviere J, Flynn T. 2007. Several methods to investigate relative attribute impact in stated preference experiments. *Social Science and Medicine* **64**: 1738-1753.

Lancsar E, Savage E. 2004. Deriving welfare measures from discrete choice experiments: Inconsistency between current methods and random utility and welfare theory. *Health Economics* **13**: 901-907.

McFadden D. 1981. Econometric models of probabilistic choice. In *Structural analysis of discrete data with economic applications,* Manski C, McFadden D, (eds.). MIT Press: Boston.

National Institute for Health and Clinical Excellence. *Social value judgements: Principles for the development of NICE guidance (2nd edition).* NICE: London, 2008.

Olsen JA, Richardson J, Dolan P, Menzel P. 2003. The moral relevance of personal characteristics in setting health care priorities. *Social Science and Medicine* **57**: 1163-1172.

Roberts T, Bryan S, Heginbotham C, McCallum A. 1999. Public involvement in health care priority setting: An economic perspective. *Health Expectations* **2**: 235-244.

Ryan M. 2004. Deriving welfare measures in discrete choice experiments: A comment to Lancsar and Savage (1). *Health Economics* **13**: 909-912; discussion 919-924.

Santos Silva JM. 2004. Deriving welfare measures in discrete choice experiments: A comment to Lancsar and Savage (2). *Health Economics* **13**: 913-918; discussion 919-924.

Sassi F, Archard L, Le Grand J. 2001. Equity and the economic evaluation of healthcare. *Health Technology Assessment* **5**: 1-138.

Small KA, Rosen HS. 1981. Applied welfare economics with discrete choice models. *Econometrica* **49**: 105-130.

Tsuchiya A, Dolan P. 2007. Do NHS clinicians and members of the public share the same views about reducing inequalities in health? *Social Science and Medicine* **64**: 2499-2503.

Tsuchiya A, Dolan P. 2008. Equality of what in health? Distinguishing between outcome egalitarianism and gain egalitarianism. *Health Economics*.

Viney R, Savage E. *Health care policy evaluation: Empirical analysis of the restrictions implied by Quality Adjusted Life Years*. CHERE Working Paper Series 2006/10. Sydney, 2006.

Williams A. 1997. Intergenerational equity: An exploration of the 'fair innings' argument. *Health Economics* **6**: 117-132.

Williams A. 1996. QALYS and ethics: A health economist's perspective. *Social Science and Medicine* **43**: 1795-1804.

Wiseman V, Mooney G, Berry G, Tang KC. 2003. Involving the general public in priority setting: Experiences from Australia. *Social Science and Medicine* **56**: 1001-1012.