A discrete choice experiment investigating preferences for funding drugs used to treat orphan diseases: an exploratory study

Emmanouil Mentzakis [a,*], Patricia Stefanowska [b] and Jeremiah Hurley [a,b]

[a] Department of Economics, McMaster University, 1280 Main St. West, Hamilton, Ontario, L8S 4M4, Canada
[b] Centre for Health Economics and Policy Analysis, McMaster University, 1200 Main St. West, Hamilton, Ontario, L8N 3Z5, Canada

JEL: I18, C25

Please cite as: Mentzakis, E., Stefanowska, P. and Hurley, J. 2011. A discrete choice experiment investigating preferences for funding drugs used to treat orphan diseases. *Health Economics, Policy and Law*, 6(3), 405-33.

Abstract

Policy debate about funding criteria for drugs used to treat rare, orphan diseases is gaining prominence. This study presents evidence from a discrete choice experiment using a convenience sample of university students to investigate individual preferences regarding public funding for drugs used to treat rare diseases and common diseases. This pilot study finds that: other things equal, the respondents do not prefer to have the government spend more for drugs used to treat rare diseases; that respondents are not willing to pay more per life year gained for a rare disease than a common disease; and that respondents weigh relevant attributes of the coverage decisions (e.g., costs, disease severity, treatment effectiveness) similarly for both rare and common diseases. The results confirm the importance of severity and treatment effectiveness in preferences for public funding. Though the first study of its kind, the results send a cautionary message regarding the special treatment of orphan drugs in coverage decision making.

1.0 Introduction

Orphan disease and their treatments are currently the focus of considerable policy attention. This policy attention arises because those who suffer from an orphan disease are perceived to be disadvantaged under the prevailing model of development for medical treatments, especially drugs. A number of factors inhibit the development of treatments for rare disease and access to those treatments that are developed. Orphan diseases are by definition rare (Wastfelt *et al.*, 2006): in Europe, an orphan disease is defined as serious, life-threatening and affecting fewer

* Corresponding author: Tel.: +1 905 525 9140 ext.24527;E-mail address: mentzak@mcmaster.ca

than 1 in 2,000 people (European Committee for Orphan Medicinal Products); Canada lacks an accepted definition, but the Canadian Organization for Rare Disorders (Canadian Organization for Rare Disorders,)defines a rare disease as affecting fewer than 1 person per 2,000 people; and in the United States the Orphan Drug Act (1983) defined an orphan disease as affecting fewer than 200,000 persons in the US[1] or more than 200,000 persons and the expectation that drug development costs will not be recovered from sales (Dear *et al.*, 2006).

Because the orphan diseases are rare, the pharmaceutical industry has little financial incentive to develop new medicines for them. The small market size makes the return on investment insufficient to attract private capital. Treatments that are developed face a series of hurdles making it to market and getting placed on insurance formularies. Clinical evidence of safety and efficacy is often less strong because of small patient samples in randomized clinical studies and the reliance on surrogate markers of effectiveness that are not always well-linked to final outcomes (Drummond *et al.*, 2007b). The high fixed costs of development and the small number of patients lead to high cost-per-patient (DiMasi *et al.*, 1991; Medecins Sans Frontieres, 2001). Consequently, relatively high incremental cost-effectiveness ratios and the poor value for money frequently lead to denial of coverage (Drummond *et al.*, 2007a).

Several governments (e.g., United States, Japan, the European Union) have introduced special financial incentives such as tax credits to spur the development of treatments ("orphan drugs") for rare diseases (Dear *et al.*, 2006; Cheung *et al.*, 2004; Denis *et al.*, 2009)[2]. Such incentives mitigate the industry's high risks and lower potential return on investments in treatments for rare diseases. These incentive schemes have increased numbers of requests for the Orphan Designation of drugs by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMEA) (Denis *et al.*, 2009). However, such policies are of limited value if treatments developed ultimately fail to get covered by insurers because of their high cost-effectiveness ratios. In response to this latter problem, some have proposed that funders apply a different, higher, cost-effectiveness threshold for drugs used to treat rare diseases (see, e.g., discussion in (Drummond *et al.*, 2007a)). This policy recommendation, however, is controversial.

Arguments for setting a higher cost-effectiveness threshold for orphan drugs vary, but the two most commonly invoked are rights-based arguments and the rule-of-rescue. Rights-based arguments posit that all members of society are entitled to access to a minimum amount of health care. Given this premise, rare disease sufferers have a right to a basic level of quality health care

[1] Given the current US population, the implied incidence rate is less than 1.3 per 2000 persons.

[2] Such policies require that "orphan treatments" be defined. The Orphan Designation procedure at the EMEA, states that to qualify a medicine must meet two conditions: a) the medicinal product is intended for the diagnosis, prevention or treatment of a life-threatening or chronically debilitating condition that either affects less than 1 in 2,000 individuals; or that without incentives is unlikely to generate sufficient return on investment to justify the expenditure and b) there is an absence of solution or the drug brings a significant benefit compared to the present situation (Denis *et al.*, 2009).

even if treatment does not offer the largest health gain for its cost (Hughes *et al.*, 2005). The rule-of-rescue principle asserts that society should come to the aid of those facing immediate, often life-threatening danger. In the orphan drug debate, the underlying premise of this principle is used to argue that society should not abandon the most severely ill individuals with rare diseases who need highly specialized treatment and have no other treatments available (Hughes *et al.*, 2005; Dolan and Olsen 2002).

Opponents of such a policy offer a corresponding set of arguments. Hughes et al. (Hughes *et al.*, 2005), argue that orphan diseases are not inherently life-threatening, although many are debilitating and reduce life-expectancy. McCabe et al. (2006) characterize arguments based on the rule-of-rescue as emotional reactions to identifiable individuals in catastrophic events, but that unknown patients will become identifiable in the future and hence it is an ethically invalid principle for policy-making. But perhaps the most common objection to setting higher cost-effectiveness thresholds for orphan drugs derive from the principle of maximizing the health gain achieved with society's limited health care resources (Schlander, 2008). The opportunity cost of such a policy is larger health losses among those who suffer from common, highly prevalent diseases (Dear *et al.*, 2006; Hughes *et al.*, 2005; McCabe *et al.*, 2006).

This debate has proceeded in a virtual vacuum of evidence regarding the views of the public regarding such a special status for drugs used to treat orphan diseases. We know that members of the public are, in general, willing to sacrifice a reduction in the total amount of health gain generated to achieve a more equitable distribution of health or health gains and to respond to those suffering severe ill health (Nord, 1993; Ubel *et al.*, 1998; Cookson and Dolan, 1999; Dolan *et al.*, 2005). We do not know, however, if this holds for responding to the needs of those with rare diseases. The only direct evidence on this point offers partial support at best. The UK's Citizen's Council of the National Institute for Health and Clinical Excellence (NICE) recommended that the National Health Service pay higher prices for ultra-orphan drugs (affecting fewer than 1000 people in the UK) provided that in addition to being very rare, the disease is severe, life-threatening, and there is evidence of health gain from treatment (NICE Citizens Council, 2004). So although rareness factors into their reasoning, rareness alone does not justify differential thresholds.

This paper presents the results of a pilot study that uses a discrete choice experiment to investigate individuals' preferences regarding public funding for drugs used to treat rare diseases. We view it as a pilot study because it is based on a convenience sample of individuals affiliated with a university. Because there is no universally accepted definition of an orphan disease or orphan drug, and the term carries strong normative overtones, we avoided use of the term "orphan drug" in the experiment and instead focused on the underlying characteristics associated with orphan diseases. We explicitly include a condition's rarity and severity as attributes in our choice scenarios as these characteristics are often emphasized in defining orphan

diseases. Further, the description of the context for the choice scenarios implies that there are no treatments available other than the drugs under consideration, as this feature also figures in debates about orphan drugs.

The study investigated three specific questions: (a) other things equal, are individuals willing to have the government pay more for drugs used to treat rare diseases than drugs used to treat common diseases; (b) other things equal, are individuals willing to have government pay more per life-year gained for a rare disease than for a common disease; and (c) in making recommendations regarding public coverage, do individual place the same relative weights on attributes across rare and common diseases?

2.0 Methods

The study employs a discrete-choice methodology (DCE) to investigate individual's preferences for funding drugs. DCEs are based on the idea that people derive utility not from a good *per se* but from the underlying attributes of the good (Lancaster, 1966). As such, a DCE presents respondents with a series of hypothetical choices that describe the choice alternatives by their underlying attributes and ask respondent which alternative they prefer. In our study, for instance, the attributes of the drug alternatives include the nature of the disease treated by a drug, a drug's effectiveness, the cost of drug treatment (more on the attributes below). The values of the attributes vary across choice scenarios, and by observing the choices people make it is possible to infer their preferences over the attributes of the goods under study. DCEs are commonly used to identify people's preferences in a variety of non-market situations/services/commodities (Bateman et al. 2002; Hensher, Rose, and Greene 2005; Louviere, Hensher, and Swait 2000) and have been widely utilized in health economics (Gerard *et al.*, 2003; King *et al.*, 2007; Ryan *et al.*, 2008; Lancsar and Louviere, 2008).

2.1 Scenario for the Discrete-Choice Experiment

To put our experiment into context, respondents were told that specialised committees meet regularly to consider adding new drugs to public drug program formularies. However, the large number of potential drugs, the limited budgets and the high costs of such programs makes public funding of prescription drugs a challenge. Participants were asked to imagine that they were a member of a government committee in the province of Ontario, Canada that makes decisions regarding drugs to be listed on the drug formulary for the province's public drug plan (see Appendix 1 for the exact description). They were told that the drug budget is limited and there are more drugs available than can be funded within the budget, so choices must be made regarding which drugs to fund. They were then told that two drugs were being considered for listing on the formulary, presented with information on the two drugs and the conditions each

drug is used to treat, and then asked which drug they would prefer to have the government fund under the public plan.

To reduce the chances that subjects might inject their own (erroneous) assumptions about the situation, the description explicitly stated that all patients were of similar age (mid-40s), marital status, income, education, etc. and could expect to live for 10 years without treatment. It was also noted that, if not treated, patients with both conditions consumed the same dollar amount of miscellaneous health care services in an effort to alleviate their symptoms. Finally, it was stated that the two drugs were identical in every respect except those characteristics explicitly described and that neither drug was associated with adverse side-effects.

2.2 DCE Attributes and Attribute Levels

Potential attributes by which to describe the choice alternatives were identified by a review of the debate about coverage decisions for orphan drugs. This identified more potential attributes than could be included in the DCE. The full list of potential attributes was reduced based on two main criteria: importance in the debate about orphan drugs (judged subjectively by frequency of mention and amount of attention given to the attribute) and ability to specify the attribute in DCE experiment. The attributes included in the scenarios were as follows (Table 1):[3]

- frequency of the disease;
- cost of treating a single patient with the drug;
- total cost of funding the drug (budget impact);
- severity of the disease without the treatment; and
- impact of drug treatment on a patient's health

Other attributes considered but not included in the final DCE (to keep the design manageable) were: the level of scientific evidence of clinical effectiveness for the drugs (orphan drugs often present lower grade clinical evidence compared to evidence for common diseases), the lack of available alternative treatments (orphan drugs might treat conditions that do not offer other therapies), and the cost-effectiveness ratios (orphan drugs often exhibit high incremental cost-effectiveness ratios and have poor value for money). The exact wording of the characteristics' descriptions is provided in Appendix 2.

Frequency of the disease

[3] As discussed below, disease frequency was incorporated through labeled alternatives in the design, so Table 1 lists the four attributes associated with each of the two disease types.

Frequency of the disease treated by a drug is a primary attribute of interest. As noted, disease frequency was used as a label in the alternatives within each choice, and not as an explicit attribute. This allowed us to specify the other four attributes as alternative-specific and to test statistically whether respondents weigh the importance of an attribute differently across rare and common diseases. Frequency took on two levels: rare and common. Following the definitions of orphan diseases, we defined a rare disease as one with an incidence rate of less than 1 case per 2000 people; correspondingly, a common disease was defined as having an incidence rate of more than 1 case per 2000 people. To aid understanding, the information was presented for a reference population of 10 million people (the approximate population of Ontario), with rare diseases having an annual incidence of fewer than 5000 cases and common diseases having an annual incidence of more than 5000 cases.

Cost of treating a single patient

Cost-per-patient was included as an indicator of the costliness of the drug treatment, and its inclusion allowed us to identify respondents' views regarding the amount the government should be willing to pay at the margin for a drug treatment. To eliminate any potential confusion, the description emphasized that the full cost of treating a single patient occurred over a three-month period and that no other costs were incurred after this treatment period. Cost-per-patient took on seven levels ranging from $1,000 to $100,000: 3 levels ($15,000, $50,000 and $100,000) occur only for rare disease alternatives, 3 levels ($1000, $5000, and $10,000) occur only for common disease alternatives, and one level ($12,000) occurs for both disease frequencies. The labeled design (described in more detail below) allows us to specify different levels for a given attribute across the two labeled alternatives, in our case, rare and common diseases. Other things equal, we expected that subjects to prefer that government fund drugs with lower cost-per-patient.

Total cost of funding the drug program

Formulary committees commonly consider not only cost-per-patient but also the total budget impact. This distinction can be particularly important for rare diseases, which can have a very high cost-per-patient but small budget impact because so few people have the disease. Total budget impact took on seven levels ranging from $5 million to $200 million: once again, 3 levels ($5 million, $10 million and $20 million) occur only for rare diseases, 3 levels ($150 million, $150 million and $200 million) occur only for common diseases, and one level ($50 million) occurs for both disease frequencies. Other things equal, we expected subjects to prefer that government fund drugs with lower total budget impact.

Severity of the disease without treatment

Severity of disease is consistently identified as a factor that individuals consider important for resource allocation in health care (Dolan and Olsen 2002). In the experiment, severity of disease if not treated could take on two values: serious and moderate impact. Severity was described in terms of the impact of the disease on a patient's quality of life and on the patient's self-assessed health status. The quality of life descriptions emphasized a patient's functioning with respect to their mobility, activities of daily living, and pain levels, and were based on health state levels of the EQ-5D classification system corresponding to the Mobility, Usual Activities, and Pain/discomfort (Dolan, 1997). "Serious" severity corresponded to a self-assessed health rating of "poor", while "moderate" severity corresponded to a self-assessed health rating of "good" [4]. Although in some instances members of the public do not prioritize those who are severely ill (Dolan *et al.*, 2008; Donaldson *et al.*, 2008), other things equal, and considering that all drugs considered return patients to full health, we expected respondents to prefer that the government fund drug that treat those with a seriously severe condition rather than those with a moderately severe condition.

Impact of drug treatment on a patient's health

The last attribute was the health gain due to the drug treatment. The health gain was specified in terms of life-years gained as a result of treatment. This attribute took on four levels ranging from 1 to 15 life years gained. In addition, respondents were told that, regardless of the baseline severity, each drug would return the patient to excellent health-related quality of life for their remaining lifetime. Other things equal, we expected respondents to prefer that the government funds drugs that produce a larger number of life-years gained.

Interactions Among Attributes

The design allowed for interactions among attributes. In choice experiments interaction effects are expected to account for a small portion of the variation (between 5 and 15 percent of the variance) and hence selected two-way interactions are normally sufficient (Hensher *et al.*, 2005). The design accounted for three interactions: that between cost-per-patient and total cost, between cost-per-patient and severity, and between severity and impact of treatment on health.

2.3 DCE design

Dependence among a subset of Attributes

[4] Using the EQ-5D scoring function from 0.0 (dead) to 1.0 (perfect health) (Dolan *et al.*, 1995), and assuming that individuals face no problems with their self-care or anxiety/depression, we get the following utilities: Serious (EQ5D=21231): 1-(0.069+0.036+0.386+.269) = 0.24; Moderate (EQ5D=11121): 1-(0.123) = 0.877. The health scores enabled us to relate our health states to utility values, but they were not presented to the respondents, who saw only the verbal descriptions.

The design of the choice experiment was complicated by the fact that three of the attributes are linearly dependent. By definition, total budget impact (or total cost, TC) equals the product of the disease frequency (F) and cost-per-patient (CP): TC=F*CP. Hence, assigning values to two of them automatically determines the third. Yet we judged it important to specify explicit quantitative values for cost attributes rather than qualitative categories such as "high cost" vs. "low cost" or "good value for money" vs. "poor value for money." To resolve this problem, we varied independently the levels of the two cost attributes and left the disease frequency (incidence rates) to be the determined attribute that varied in the background within defined bounds consistent with a rare disease or a common disease. That is, the two cost attribute levels were chosen in such way that the implied incidence rates were lower than 1 in 2000 cases for rare alternatives and greater than 1 in 2000 for common alternatives. For example, if the total cost is assumed to be $50 million, a cost-per-patient of less than $25,000 is consistent only with a common disease (frequency of more than 5000, or an incidence rate of more than 1 in 2000 in a population of 10 million) and a cost-per-patient of $25,000 or more is consistent only with a rare disease (frequency of 5000 or less). Combinations of total cost and cost-per-patient correspond accordingly with incidence rates of less than 1 in 2000 people for a rare disease and greater than 1 in 2000 people for a common disease. Respondents never saw the implied incidence rates; they simply saw the labels "rare" or "common" disease.

We employed a labelled, forced-choice experimental design. Unlike generic experiments, labelled experiments brand each alternative, which subsequently carries information and meaning that is likely to influence the choice outcomes. Moreover, such designs allow for different sets of attribute levels across the alternatives. Hence, for every decision, respondents faced a choice between: a drug used to treat a rare disease with specified attribute levels for each of cost-per-patient, total budget impact, severity of disease and life-years gained by treatment; and a drug used to treat a common disease with correspondingly specified attribute levels.

Forced experiments constrain respondents to express a preference (i.e., make a trade-off among attributes) even when both alternatives are unattractive. Hensher *et al.* (2005) argue that such a design is preferred when the objective of the study is to examine "the impact of the relationships different attribute levels have upon choice" (p. 176), such as is the case in our setting.

2.4 Experimental design

A full-factorial, labelled design with three four-level attributes and one two-level attributes generates 16,384 possible combinations ($L^{MA} = 4^{2*3} * 2^2$=16,384), hence a fractional factorial design was used. Allowing for two-way interactions a D-efficient (D-efficiency = 0.817) fractional factorial design was produced with 64 pair-wise choices (Kuhfeld, 2005; Zwerina *et*

al., 1996), which we blocked into 4 blocks of 16 choices each. All aspects of the experimental design were performed using SAS 9.1.3 built-in capabilities (Kuhfeld, 2005).

2.5 Econometric methods

The utility an individual derives from choosing an alternative is assumed to comprise two components, a deterministic component and a stochastic component. Assuming an additive deterministic component to utility, $V_{iq} = \sum_{k=1}^{K} \beta_k X_{ikq}$, and a stochastic component to utility, ε_{iq}, the utility of an individual *q* choosing alternative *i* is

$$U_{iq} = V_{iq} + \varepsilon_{iq} = \sum_{k=1}^{K} \beta_k X_{ikq} + \varepsilon_{iq} \qquad \text{(eq. 1)}$$

where *i* denotes the choice alternative, k denote the attributes and q denotes an individual, and β_k are the utility parameters to be estimated. β_k are assumed to be homogeneous across the population.

Conditional Logit versus Latent Class Models

Assuming ε_{iq} to be independent and identically distributed (iid) extreme value type I (EV1), $F(\varepsilon) = \exp(-\exp(-\varepsilon))$ gives rise to the McFadden's (1974) conditional logit (CL), for which the probability that individual *q* chooses alternative *i* from a among a set of J alternatives is:

$$P_{iq} = \frac{\exp(V_{iq})}{\sum_{j=1}^{J} \exp(V_{jq})} \qquad \text{(eq. 2)}$$

The conditional logit model has a number of attractive features and is the standard approach to analyzing data from choice experiments, but it does impose some restrictive assumptions that often fail to hold (Hensher *et al.*, 2005). In particular, its assumption of independence of irrelevant alternatives (IIA), its failure to incorporate preferences heterogeneity in the utility parameters and its inability to account for the panel structure of data has led researchers to identify models with more flexible structures. A commonly used alternative is the mixed logit model; and more recently, some have argued that the semi-parametric latent-class model (LCM) often performs as well or better than the mixed logit (Greene and Hensher, 2003; Hole, 2008).

In the LCM, parameter heterogeneity across individuals is modeled with a discrete distribution over a set of classes. "Individuals are implicitly sorted into a set of *C* classes, but which class contains any particular individual, whether known or not to that individual, is unknown to the analyst" (Greene and Hensher, 2003) (p.682). The IIA is imposed only within classes and not on

the observed unconditional probabilities. The probability that individual *q* chooses alternative *i* in choice set *t* conditional on falling within class *c* is

$$P_{iqt|c} = \frac{\exp(X_{iqt}\beta_c)}{\sum_{j=1}^{J}\exp(X_{jqt}\beta_c)} \quad \text{(eq. 3)}$$

Following Greene and Hensher (2003) let y_{qt} denote a specific choice made such that $P_{qt|c} = Prob(y_{qt} = i \mid class = c)$. Hence, given a specific class assignment

$$P_{q|c} = \prod_{t=1}^{T} P_{qt|c} \quad \text{(eq. 4)}$$

Additionally, let H_{qc} be the probability that individual *q* falls in class *c*

$$H_{qc} = \frac{\exp(z_q\theta_c)}{\sum_{c=1}^{C}\exp(z_q\theta_c)} \quad \text{(eq. 5)}$$

where z_q is a set of variables that characterize the probabilities for class membership. Following from eq. 4 and eq. 5, for *c* classes the likelihood for individual *q* is

$$P_q = \sum_{c=1}^{C} H_{qc} \cdot P_{q|c} \quad \text{(eq. 6)}$$

The number of latent classes is not determined endogenously but is determined *a priori*, based on the performance of alternative models with respect to information criteria such as the Akaike (AIC), the Bayesian (BIC) and the Hannan-Quinn (HQIC) (Hole, 2008; Swait and Adamowicz, 2001; Hannan and Quinn, 1979). We estimate and compare both conditional logit and latent-class models.

Variable Specification

In addition to the choice attributes, our specification of V_{iqt} includes three two-way interactions (total cost · cost-per-patient; cost-per-patient · severity; severity · life-years-gained) and information on cut-offs as described in Section 2.6. As explained below, such modifications allow for the relaxation of restrictive assumptions regarding the linearity of the attribute effects. Because of the labeled design, all attributes can be specified as generic or alternative-specific; a Wald test for the equality of attribute coefficients across rare and common alternatives can identify if alternative-specific attributes are appropriate.

We present results in three ways: the estimated model coefficients, marginal rates of substitution among attributes (Bennett and Adamowicz, 2001) (with standard errors calculated using the delta method), and changes in the predicted probability of choosing a choice alternative caused by changes in the choice attributes. We are particularly interested in marginal rates of substitution between attributes and cost, which conveys marginal willingness to pay for an attribute. The marginal willingness to pay values are computed for changes evaluated at the means of the attributes; for the LCM they are (along with the estimated coefficients), class specific. The changes in probabilities are computed for unitary changes at the mean of the regressors for total budget, cost-per-patient and life-years gained and for discrete changes for frequency and severity, while keeping the rest of the attributes to their sample mean values. All estimation and calculations were performed using Nlogit 4.0.

2.6 Allowing for discontinuities in Marginal Utility by Incorporating Decision Cut-offs

By definition, discrete choice experiments are based on trades-offs between attributes. More of one attribute is assumed to compensate for less of another (Louviere, Hensher, and Swait, 2000). However, in making choices individuals often employ decision heuristics that violate such compensatory behavior. Such heuristics can include, elimination-by-aspects (Tversky, 1972) and conjunctive rules (Dawes, 1964) in which subjects follow cut-offs when making choices. A cut-off is a decision rule that sets limits beyond which the subject would never choose an alternative (e.g. a 'rule' that one would never buy a house without a swimming pool or they would never pay more than $X for a specific product). However, even when individual use such cut-off rules, the evidence shows that they often do not follow them strictly (Huber and Klein, 1991; Swait, 2001). In a sense, rather use hard cut-offs individuals seem to employ soft cut-offs, or thresholds, around which the marginal utility of an attribute level varies. This possibility is important because standard discrete-choice experiments assume no such discontinuities in the marginal utility of attributes.

The difficult nature of the decision problems we presented subjects regarding drug coverage suggested to us that they might employ non-compensatory decision heuristics. We therefore included in our study a component that would enable us to test for discontinuous utility with respect to attributes. In implementing this we followed Swait (2001), who proposed a penalising utility function that allows for cut-off violation and approximates/simulates a number of non-compensatory behaviours. In this framework, decision cut-offs are not "hard" in the sense that a person never violates them; rather, it assumes that subjects suffer a greater loss in utility at the margin the further one is from the threshold (i.e., there is a utility penalty for such a choice).

Implementing this approach required that we collect, for each attribute in the study, information on the value of each respondent's cut-offs (below or above which they would "never" choose an

alternative). For example, if a subject agreed that "the government should not fund drug treatments that extend life fewer than 5 years", this would provide information regarding the lower cut-off (or threshold) the subject placed on the attribute life-years-gained.

Following Swait (2001), for choice *i* and *k* attributes, allowing for cut-offs modifies the deterministic component of eq. 1, as follows

$$U_i = \sum_k \beta_k X_{ik} + \lambda_k \sum_k w_k \cdot \max(0, c_k - X_{ik}) + \tau_k \sum_k v_k \cdot \max(0, X_{ik} - d_k) \qquad \text{(eq. 7)}$$

Where c_k and d_k are the lower and upper cut-offs as stated by the respondent, $\max(0, c_k - X_{ik})$ identifies the magnitude of the violation of the lower bound cut-off and $\max(0, X_{ik} - d_k)$ identifies the corresponding violation for the upper bound cut-off. For an attribute with a negative marginal utility (e.g., cost), let the initial negative slope be β_k (see Figure 1). If a subject were to choose an alternative with a cost over a specific upper threshold "d", marginal utility falls at a higher rate ($\beta_k + v_k$); if they choose an alternative with cost below "c", marginal utility falls at a lower rate ($\beta_k + w_k$). Marginal utility displays kinks, where "c" and "d" are the lower and upper cut-offs. Our specification does not include both upper and lower threshold for an attribute, but it does include upper cut-offs for some attributes and a lower cut-off for others, so eq. 7 above includes two indicator variables, λ_k and τ_k, where λ_k takes on a value of 1 if the attribute is hypothesized to have a lower cut-off and 0 otherwise, and τ_k takes on value of 1 if the attribute is hypothesized to have an upper cut-off and 0 otherwise.

As, Swait (2001) further shows, cut-off violation is possible even for binary indicators, where a violation becomes itself a dummy variable indicating a case where the level of the attribute differs from that stated in the cut-off elicitation exercise. The subjects completed the cut-off elicitation questions before completing the choice experiment. The elicitation method involved asking subjects a series of questions formulated as "I would never pay more than $X to fund this drug" (see Appendix 3).

The cut-off information is integrated into the analysis by including in the regression the second and third terms of eq. 1, with their coefficients estimated along with the rest of the parameters. The magnitude and statistical significance of the associated coefficient estimates provides a test of the importance of discontinuities in the subjects' decisions (compared to a model that assumes cut-offs are not present) (Danielis and Marcucci, 2007). Non-significant coefficients on the cut-off variables imply that such cut-off values play no role in decision-making.

2.7 Survey Development and Administration

Development of the survey instrument was refined by two pre-tests. The first pre-test was conducted among a convenience sample of colleagues, research staff and graduate students early in the development of the instrument. This pre-test focused on basic aspects of the survey design such as the instructions, design of the choice scenario, and specification of the attributes. The second pre-test was conducted among a sample of 50 individuals drawn from the subject pool and focused on final refinement of the survey, clarity and understanding by respondents, and testing the procedures for administering the survey. As part of this pre-test respondents also completed open-ended questions regarding the clarity and difficulty of the content and length of time required to complete the survey. In addition, five in-depth interviews with a random sample of the pre-test subjects provided further insight into points of ambiguity or other problems. Revisions were made in light of the feedback.

For the main survey, a random sample of individuals drawn from a subject database maintained by the McMaster University Experimental Economics Laboratory and invited to participate in the study. Following recommendations on sample size (Hensher *et al.*, 2005; Lancsar and Louviere, 2008), 20 to 50 participants per block was deemed adequate for robust estimation, suggesting a target a sample of 200 individuals. The vast majority of subjects were students. Past experimental valuation studies (Maguire *et al.*, 2003; Depositario *et al.*, 2009) have concluded that the views of students often closely represent those of the broader community of non-students. However, given the use of a convenience sample of individuals from the university, we view this study as a pilot that can produce initial evidence and inform the design of subsequent studies in the general community. All participants were compensated $8 for their participation.

The full survey was administered electronically in the McMaster University Experimental Economics Laboratory. Ethics approval for the study was obtained from the Hamilton Health Sciences/McMaster University Research Ethics Board.

3.0 Results

3.1 Sample descriptive statistics

213 respondents completed the survey. The sample characteristics are as follows (Table 2): 59% was female; mean age was approximately 22 years; 80% reported excellent or very good health status; about 30% had a part-time job; for 6.5% and 17.4% the father or the mother, respectively, was unemployed; for approximately 87%, their parents owned their house; and 32% financed their education at least partly through a registered education savings plan (RESP), which are generally used by those with above-average income.

3.2 Multivariate Results

Model Selection

Information criteria (AIC, BIC, HQIC) indicated that the 2-class LCM performed better than the traditional conditional logit[5] (CL: AIC = 1.026, BIC= 1.048, HQIC = 1.033; LCM: AIC = 0.97, BIC=1.047, HQIC = 1.003). Within the LCM, specifying more than 2 classes often resulted in convergence problems and singularities in the variance matrices. However, in cases in which convergence was achieved, the specification with 2 classes was preferred to the specification with more than two classes. A test of the joint statistical significance of the alternative-specific attributes indicated that a generic model was preferred (χ^2 =15.058; p-value = 0.37). We therefore present results from the 2-class LCM with generic attribute-coefficients. Finally, an LR-test (χ^2 = 51.1; p-value = 0.000) implied the joint significance of the cut-offs and, hence they were kept in the model.

Coefficient Estimates

Table 3(a) presents the results of the logit model regarding the probability each individual falls into each of the two latent classes, with assignment being based on a set of personal characteristics. Among the characteristics included in the model only sex, father's employment status and whether the respondent's university education is at least partly financed by a registered education savings plan are statistically significant predictors of class membership. The table presents the logit coefficients, so the results imply that the odds that a female is in class 1 are 0.42 times those of than a male; the odds that a subject whose father is unemployed is in class 1 are 0.08 those of a person whose father is employed; and the odds that a subject who finances their education at least partly through a RESP is in class 1 are 2.81 times those who do not.[6] The two classes are of approximately equal size with the average probability that a respondent falls into class 1 being 46% and the average probability for class 2 being 54%.

Table 3(b) presents the attribute coefficients for each class. We discuss the estimates separately for each class and then compare them. For class 1, the significant common disease intercept indicates that, all else equal, respondents preferred that the government fund drugs for common diseases. The coefficients on both total budget impact and cost-per-patient are not statistically significant for either common or rare diseases, implying that neither cost attribute influenced preferences over funding for a drug. Neither are the interaction terms involving these two cost attributes (TC*CP and CP*SEV) statistically significant. The coefficients for baseline disease severity and for life years gained by treatment are significant and positive. As expected,

[5] For completeness the LCM was also compared to a MXL and both fit the data equally well. We present the LCM results because the LCM is a more parsimonious model and the estimates are more easily interpreted with respect to the relevant policy issues.

[6] Odds ratios are obtained by exponentiating the coefficients, i.e. exp(-0.8773)=0.42; exp(-2.5895)=0.08; exp(1.0317)=2.81.

respondents prefer that government fund a drug to treat a serious condition rather than a condition of moderate severity, and that it fund a drug that produces more life years gained. Table 4 presents the impact of a one-unit change (or discrete change) in each attribute on the probability that subjects prefer funding a drug. The change from rare to common disease increases this probability by about 30 percentage points, while the corresponding effect for the two cost attributes is not statistically different from zero. For severity, the probability that a subject prefers to have government fund a drug used to treat severe condition was 22 percentage points higher than for a moderate condition. For life-years gained by treatment, an increase of 1 life-year gained increases by 4.5 percentage points the probability that a subject prefers to have government fund a drug.

For class 2, though the trend again favours a common disease, the common disease intercept is smaller and not statistically significant, indicating that frequency of disease exerted little influence on subjects choices. The coefficient estimates on both total budget impact (TC) and cost-per-patient (CP) are negative and statistically significant. It appears, therefore, that for class 2 costs influence choice in the expected direction (greater cost reduces the probability of funding a drug). The coefficient on the interaction term TC*CP is statistically significant and possesses an unexpected positive sign. However, the aggregate effect (TC+CP+TC*CP) is negative over values of CP and TC that appear in our design.[7] The implied magnitudes of these cost effects are as follows: a $1 million dollar increase in the total budget impact is predicted to decrease the probability that a drug is chosen by a 0.07 percentage points; for cost-per-patient, an $1000 increase in cost-per-patient is predicted to decrease the probability that a drug is chosen by 0.12 percentage points. As we saw for class 1, the coefficient estimates for both baseline severity of illness and life-years gained are positive and statistically significant, and the estimates again imply large effects on choice. The probability that a drug is chosen for government funding is 29.5 percentage points higher if it is used to treat a severe condition than if it is used to treat a moderate condition. And an increase in 1 life-year gained raises the probability that a drug is chosen for government funding by 4.9 percentage points. The interactions term CP*SEV has an unexpected positive sign, implying that for cases involving a serious disease an increase in CP makes the alternative more attractive. As expected, the interaction effect is positive for severity and life years (SEV*LYG).

The coefficients estimates associated with the cut-off analysis are presented at the bottom of Table 3. None of the estimates approach statistical significance for class 1, while for class 2 there is evidence of discontinuity with respect to two attributes: disease severity and life-years gained by treatment. Hence, for those in the sample that stated an *a priori* intention to not fund anything with a moderate severity but subsequently violated it, the estimate implies that the "penalty" is -0.191 "utils": the coefficient for choosing a moderate condition is 0 (i.e., baseline of a binary

[7] The aggregate effect becomes positive only when the three highest TC values (100, 150, 200) are interacted with the highest CP value (100), which does not appear in our design (see attribute levels definition, Table 1).

indicator), hence, the penalty is: 0 + (-0.191)). Similarly, for those stating an intention to not fund anything that extended life less than 5 years, choosing a drug that generates fewer than 5 additional life-years has a negative impact on utility. The magnitude of the penalty is larger than that the effect of the attribute itself and implies that for drugs that produce fewer than 5 additional life-years, the utility obtained from life years is negative (0.1823 - 0.5335 = -0.3512).

Marginal rates of substitution

Tables 5 and 6 present marginal rates of substitution (MRS) among selected attributes. The calculated MRSs are meaningful only when both attribute coefficients are statistically significant. Hence for class 1 (Table 5), we present MRSs only with respect to severity and life years gained by treatment. On average, individuals of class 1 are willing to forgo 4.6 life years to fund a drug that treats a serious condition rather than a moderate one. This implies that individuals are equally willing to have government fund a drug that treats a serious condition as a drug that treats a moderate condition and produces an additional 4.6 life years gained for recipients.

Individuals in the second class (Table 6) are willing to forgo slightly more life-years gained (7.4) to treat a serious condition rather than a moderate one. The significance of the cost attributes in class 2 allows for the calculation of monetary marginal willingness-to-pay. Individuals are willing to have government spend an extra $174,640 per-patient to fund a drug used to treat a serious condition compared to a drug used to treat a moderate condition. They are willing to have government incur an additional total cost of $190 million to treat a serious rather than moderate condition.

Finally, all else equal, individuals are willing to have the government spend an additional $26,160 cost-per-patient for an extra life-year gained by treatment and an additional $27.44 million in total to fund a drug that provides an extra life year gained for all those who receive treatment.

4.0 Discussion

This exploratory study is a first attempt to present empirical evidence regarding individual preferences about funding for drugs used to treat orphan diseases. The frequency of a disease in a population is only one feature of a disease relevant to funding coverage decisions, making discrete-choice methodology well-suited for investigating how and to what extent disease frequency influences people's judgments regarding public funding for a drug.

Our study was designed to answer three specific questions with regard to coverage decisions: (a) other things equal, are individuals willing to have the government pay more for drugs used to

treat rare diseases than drugs used to treat common diseases; (b) other things equal, are individuals willing to have government pay more per life-year gained for a rare disease than for a common disease; and (c) in making recommendations regarding public coverage, do individual place the same relative weights on attributes across rare and common diseases.

Our results indicate that the answer to the first question is no: other things equal, a subset of people (class 1) actually appear willing to have government pay less for drugs used to treat rare diseases; however, another subset (class 2) makes no distinction based on frequency of disease.

For the second question our results indicate the answer to be also no: people do not appear willing to pay more per life-year gained for those who suffer from a rare disease than those who suffer from a common disease. Statistical tests rejected the alternative-specific models in favour of a single generic specification, which implies that individuals do not weigh attributes differently across common and rare diseases when making funding decisions. Our results, however, do indicate heterogeneous preferences between the two latent classes. For those who fall in latent class 1, costs do not exert a meaningful influence on choice; for those who fall in latent class 2 and for whom costs do exert an important influence on decisions, the willingness-to-pay for an additional life year gained by treatment did not differ between common and rare diseases.

Finally, the results indicate that the answer to the third question is no: respondents did not weight attributes in a meaningfully different way across common and rare diseases. The Wald test could not reject the null hypothesis of no difference, so that coefficient estimates are not statistically different from each other.

We do find large effects for those attributes that the literature on priority-setting suggests we should: severity of disease and treatment effectiveness. For both latent classes, the influence of these two attributes outweighs all other influences. These findings are also consistent with the findings of NICE's Citizens Council — the only other evidence currently available regarding the public's views on coverage for drugs used to treat orphan diseases — that rareness itself does not justify special consideration, but rather that severity, established evidence of effectiveness and the life-threatening character of the disease weighed more heavily (NICE Citizens Council, 2004). Our findings regarding the lack of emphasis on costs for those in class 1 are consistent with a Canadian study that used a citizen jury to investigate public views on criteria for priority-setting for health technology assessment. Members of the citizen jury argued that cost should not be considered during priority-setting for HTA but that high incidence, quality of life and potential to extend life should be among the top criteria (Menon and Stafinski, 2008). In particular, "of greatest importance was the 'potential to benefit a number of people' reflecting the jury's view that technologies for highly incident/prevalent conditions within a population should be given higher priority over less prevalent ones.

The differences in choice behavior across the two classes may reflect the broader divisions in the debate about coverage decision-making. There is no such thing as a single "view"; rather, there is systematic heterogeneity that the LCM was able to identify. Specifically, one approach to coverage decision-making (Class 1) put little weight on cost considerations, focusing almost exclusively on severity and treatment effect; in contrast, the second (Class 2) considered costs in addition to severity and treatment effect. Note that an individual's class cannot be identified (i.e. latent) but overall, across the whole sample, the two tendencies are almost of equal frequency. Hence, in debates about coverage decision-making, one would expect differing positions, one of which emphasizes the importance of considering costs and one who which downplays such considerations. Interestingly, our results also suggest that the decisions of males and those of higher socio-economic status (father employed and finances education through registered savings accounts) have a stronger tendency to downplay considerations of costs. The ability to accommodate such heterogeneity is a strength of the LCM and the likely reason why both in this and previous studies (Greene and Hensher, 2003; Hole, 2008) it performed better than the more commonly applied CL model.

Our results find some evidence of discontinuity for two attributes in Class 2, indicating that the marginal utilities of severity and life-years gained are not constant over the whole range of values presented. However, the forced-choice nature of our design may have compromised our ability to accurately identify such effects and exaggerated their presence. For instance, an individual may have faced two alternatives, both of which violated a decision cut-off (e.g., life-years gained for each alternative was below the respondent's self-declared cut-off), but still was forced to choose one as the preferred alternative. Even with such potential exaggeration, raw averages present little evidence of cut-off violation (Table 2). It should be noted that our main results are not sensitive to the inclusion or exclusion of the cut-off variables from the model.

The assumption that treatment restored individuals to full health resulted in some loss of realism for the choice scenario, but was done for two primary reasons. First, it created a cognitively simpler choice task for the respondents. Pre-testing revealed that explicitly varying baseline severity, number of life-years gained and post-treatment quality of life was cognitively much more difficult to evaluate. Second, including a separate attribute for post-treatment quality of life would create saliency problems and would require imposing design restrictions with respect to severity. For example, a drug used to treat a severe disease that resulted in a higher quality of life than a drug used to treat a moderate disease could potentially raise questions about the effectiveness of the drugs, which was outside the scope of this study.

Further, our statement in the scenario description that all patients were of a similar age (mid-40s) was included ensure that respondents used the same reference point for all alternatives when

making choices. There is evidence that individual preferences regarding health can vary with the age of the patient (Dolan *et al.*, 2005)), but it potentially limits the generalizability of our results to patient populations of a different age.

Overall, our results indicate that the participants (in this case, represented by a university-affiliated sample from Ontario, Canada) do not support differential consideration of orphan diseases for coverage decision-making. They therefore send a cautionary message about implementing special coverage rules for drugs to treat orphan diseases. Two considerations, however, should determine the role of this evidence in the debate. First, even assuming we have accurately elicited the views of the public, such preferences are not necessarily determinative in resolving difficult ethical problems such as funding treatments for orphan diseases. The resolution of such issues normally requires consideration of both sound ethical reasoning from principles and the preferences and attitudes of members of society. Second, as the first empirical results on this question, these findings should be seen as tentative and subject to further research both in different populations (i.e. generalizability of student samples is hard to advocate) and using modified designs (i.e. more comprehensive utility functions e.g. degree of rareness) to validate any conclusions. Although it is reasonable to observe differences in views across populations, such research can help identify potential framing effects associated with any single study and those aspects of the results that are robust across studies.

Acknowledgements

The authors are grateful to two anonymous reviewers for their valuable suggestions. Also we would like to thank Mita Giacomini and Jean-Eric Tarride for helpful comments and advice. We thank colleagues in the NET group at McMaster University for useful feedback regarding the design of the experiment and the analysis of the data, and we thank David Cameron and Aleksandra Gajic for assistance administering the survey. This research was funded by the Canadian Institutes of Health Research (Grant # 76670). We also acknowledge funding from the Ontario Ministry of Health and Long-term Care to the Centre for Health Economics and Policy Analysis, McMaster University. The views expressed are those of the authors alone.

References

Bateman, I., R.T. Carson, B. Day, W.M. Hanemann, N. Hanley, T. Hett, A. Jones, *et al.* (2002), *Economic valuation with stated preferences techniques: A manual*. Cheltenham, UK: Edward Elgar.

Bennett, J., and W. Adamowicz. (2001). 'Some fundamentals of environmental choice modelling', in J. Bennett, R. Blamey (eds.), *The choice modelling approach to environmental valuation.*, Northampton: Edward Elgar.

Canadian Organization for Rare Disorders. [cited September 5 2008]. Available from http://www.cord.ca/.

Cheung, R.Y., J.C. Cohen, and P. Illingworth (2004), 'Orphan drug policies: Implications for the united states, canada, and developing countries.', *Health Law Journal*, 12: 183-200.

Cookson, R., and P. Dolan (1999), 'Public views on health care rationing: A group discussion study', *Health Policy*, 49(1-2): 63-74.

Danielis, R., and E. Marcucci (2007), 'Attribute cut-offs in freight service selection', *Transportation Research Part E: Logistics and Transportation Review*, 43(5): 506-15.

Dawes, R.M. (1964), 'Social selection based on multidimensional criteria', *Journal of Abnormal Psychology*, 68: 104-9.

Dear, J.W., P. Lilitkarntakul, and D.J. Webb (2006), 'Are rare diseases still orphans or happily adopted? the challenges of developing and using orphan medicinal products.', *British Journal of Clinical Pharmacology*, 62(3): 264-71.

Denis, A., S. Simoens, C. Fostier, L. Mergaert, and I. Cleemput. (2009), 'Policies for orphan diseases and orphan drugs', KCE reports 112C, Belgian Health Care Knowledge Centre, .

Depositario, D.P.T., R.M. Nayga Jr., X. Wu, and T.P. Laude (2009), 'Should students be used as subjects in experimental auctions?', *Economics Letters*, 102(2): 122-4.

DiMasi, J.A., R.W. Hansen, H.G. Grabowski, and L. Lasagna (1991), 'Cost of innovation in the pharmaceutical industry', *Journal of Health Economics*, 10(2): 107-42.

Dolan, P., R. Edlin, and A. Tsuchiya. (2008), 'The relative societal value of health gains to different beneficiaries', RM03/JH11.

Dolan, P., and J.A. Olsen. (2002), *Distributing health care: Economic and ethical issue*s. New York: Oxford University Press Inc.

Dolan, P. (1997), 'Modeling valuations for EuroQol health states', *Medical Care*, 35(11): 1095-108.

Dolan, P., R. Shaw, A. Tsuchiya, and A. Williams (2005), 'QALY maximisation and people's preferences: A methodological review of the literature', *Health Economics*, 14(2): 197-208.

Donaldson, C., R. Baker, I. Bateman, M. Jones-Lee, E. Lancsar, G. Loomes, H. Maso*n, et al*. (2008), 'Weighting and valuing quality adjusted life years: Preliminary results from the social value of a QALY project'. University of Birmingham, Working paper JH12 .

Drummond, M.F., D.A. Wilson, P. Kanavos, P. Ubel, and J. Rovira (2007a), 'Assessing the economic challenges posed by orphan drugs', *International Journal of Technology Assessment in Health Care*, 23(01): 36.

———. (2007b), 'Assessing the economic challenges posed by orphan drugs: A response to McCabe et al.', *International Journal of Technology Assessment in Health Care*, 23(03): 401-4.

Gerard, K., M. Shanahan, and J. Louviere (2003), 'Using stated preference discrete choice modelling to inform health care decision-making: A pilot study of breast screening participation', *Applied Economics*, 35(9): 1073-85.

Greene, W.H., and D.A. Hensher (2003), 'A latent class model for discrete choice analysis: Contrasts with mixed logit', *Transportation Research Part B: Methodological*, 37(8): 681-98.

Hannan, E.J., and B.G. Quinn (1979), 'The determination of the order of an autoregression', *Journal of the Royal Statistical Society.Series B (Methodological)*, 41(2): 190-5.

Hensher, D.A., J.M. Rose, and W.H. Greene. (2005), *Applied choice analysis: A prime*r. Cambridge, UK: Cambridge University Press.

Hole, A.R. (2008), 'Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment', *Journal of Health Economics*, 27(4): 1078-94.

Huber, J., and N.M. Klein (1991), 'Adapting cutoffs to the choice environment: The effects of attribute correlation and reliability', *The Journal of Consumer Research*, 18(3): 346-57.

Hughes, D.A., B. Tunnage, and S.T. Yeo (2005), 'Drugs for exceptionally rare diseases: Do they deserve special status for funding?[see comment]', *Qjm*, 98(11): 829-36.

King, M.T., J. Hall, E. Lancsar, D. Fiebig, I. Hossain, J. Louviere, H.K. Reddel, and C.R. Jenkins (2007), 'Patient preferences for managing asthma: Results from a discrete choice experiment', *Health Economics*, 16(7): 703-17.

Kuhfeld, W.F. (2005), 'Marketing research methods in SAS', TS-722, SAS Institute Inc., Cary, NC, USA.

Lancaster, K. (1966), 'A new approach to consumer theory', *The Journal of Political Economy*, 74(2): 132-57.

Lancsar, E., and J. Louviere (2008), 'Conducting discrete choice experiments to inform healthcare decision making: A user's guide', *PharmacoEconomics*, 26(8): 661-77.

Louviere, J.J., D.A. Hensher, and J.D. Swait. (2000), *Stated choice methods: Analysis and application*s. Cambridge, UK: Cambridge University Press.

Maguire, K.B., L.O. Taylor, and S. Gurmu (2003), 'Do students behave like adults? evidence from valuation experiments', *Applied Economics Letters*, 10: 753-6.

McCabe, C., A. Tsuchiya, K. Claxton, and J. Raftery (2006), 'Orphan drugs revisited', *Qjm*, 99(5): 341-5.

McFadden, D. (1974). 'Conditional logit analysis of qualitative choice behavior', in P. Zarembka (ed.), *Frontiers in econometrics*. New York: Academic Press.

Medecins Sans Frontieres. (2001), 'Fatal imbalance. the crisis in research and development for drugs for neglected diseases', Medecins Sans Frontieres Access to Essential Medicines Campaign and the Drugs for Neglected Diseases Working Group, Geneva.

Menon, D., and T. Stafinski (2008), 'Engaging the public in priority-setting for health technology assessment: Findings from a citizens' jury', *Health Expectations*, 11(3): 282-93.

NICE Citizens Council. (2004), 'NICE citizens council report: Ultra orphan drugs', National Institute for Health and Clinical Excellence, London: .

Nord, E. (1993), 'The trade-off between severity of illness and treatment effect in cost-value analysis of health care', *Health Policy*, 24(3): 227-38.

Ryan, M., K. Gerard, and M. Amaya-Amaya, eds. 2008. *Using discrete choice experiments to value health and health car*e. The economics of non-market goods and resources, ed. I. J. Bateman. The Netherlands: Springer.

Schlander, M. (2008), 'The use of cost-effectiveness by the national institute for health and clinical excellence (NICE): No(t yet an) exemplar of a deliberative process', *Journal of Medical Ethics*, 34(7): 534-9.

Swait, J. (2001), 'A non-compensatory choice model incorporating attribute cutoffs', *Transportation Research Part B: Methodological*, 35(10): 903-28.

Swait, J., and W. Adamowicz (2001), 'The influence of task complexity on consumer choice: A latent class model of decision strategy switching', *Journal of Consumer Research*, 28(1): 135-48.

Tversky, A. (1972), 'Elimination-by-aspects: A theory of choice', *Psychology Review*, 79: 281-99.

Ubel, P.A., M.D. Spranca, M.L. Dekay, J.C. Hershey, and D.A. Asch (1998), 'Public preferences for prevention versus cure: What if an ounce of prevention is worth only an ounce of cure?', *Medical Decision Making*, 18(2): 141-8.

Wastfelt, M., B. Fadeel, and J.I. Henter (2006), 'A journey of hope: Lessons learned from studies on rare diseases and orphan drugs', *Journal of Internal Medicine*, 260(1): 1-10.

Zwerina, K., J. Huber, and W.F. Kuhfeld. (1996), 'A general method for constructing efficient choice designs', SAS Technical Papers: TS-722E.

Table 1: Attributes definitions and levels

Attributes	Levels	Coding
Rare disease alternative		
The cost of treating a single patient (CP)	$12,000	12
	$15,000	15
	$50,000	50
	$100,000	100
The total cost of funding the drug program (TC)	$5 million total cost	5
	$10 million total cost	10
	$20 million total cost	20
	$50 million total cost	50
The severity of the disease without treatment (SEV)	Serious Impact	Serious Impact=1
	Moderate Impact	Moderate Impact=0
The impact of drug treatment on a patient's health/life years gained (LYG)	15 years	15
	10 years	10
	5 years	5
	1 year	1
Common disease alternative		
The cost of treating a single patient (CP)	$1,000	1
	$5,000	5
	$10,000	10
	$12,000	12
The total cost of funding the drug program (TC)	$50 million total cost	50
	$100 million total cost	100
	$150 million total cost	150
	$200 million total cost	200
The severity of the disease without treatment (SEV)	Serious Impact	Serious Impact=1
	Moderate Impact	Moderate Impact=0
The impact of drug treatment on a patient's health/life years gained (LYG)	15 years	15
	10 years	10
	5 years	5
	1 year	1

Table 2: Sample descriptive statistics

Total # participants	*213 (100%)*
Individual characteristics	
Mean age (sd; min, max)	21.6 (4.9; 18, 60)
Sex	
Female	125 (58.7%)
Self-assessed Health Status	
Excellent/Very Good	171 (80.3%)
Know someone with chronic disease (Base = No)	
Yes-Rare disease (Dis_Rare)	35 (16.4%)
Yes-Common disease (Dis_Com)	87 (40.9%)
McMaster University Status	
Graduate	10 (4.7%)
Faculty/Staff	9 (4.3%)
Work Status	
Working part-time	63 (29.6%)
Working full-time	10 (4.7%)
Part of your university education paid with funds from a Registered Education Savings Plan (RESP)	
Yes	69 (32.4%)
Family characteristics	
Parent's housing tenure	
Owner	185 (86.9%)
Father's employment status	
Unemployed/Not applicable	18 (6.5%)
Mother's employment status	
Unemployed/Not applicable	37 (17.4%)
Stated cut-offs (% of choices that violated the cut-off)	
TC140: Cut-off violated when an individual chooses alternative with TC higher than 140mil, when initially s/he had identified it as the maximum TC they would be willing to incur.	11.4%
CP80: Cut-off violated when an individual chooses alternative with CP higher than 80 thousand, when initially s/he had identified it as the maximum CP they would be willing to incur.	4.7%
LYG5: Cut-off violated when individual chooses alternative with LY gained of less than 5 years, where initially s/he had identified it the	7.6%

minimum amount they would require.	
SevSer: Cut-off violated when individual chooses alternative with moderate severity, while initially s/he had stated that government should only fund diseases with serious impact.	19.7%

Table 3. Latent Class Model Results (2 Classes)

(a) Logistic Regression Results for Class Assignment		
Constant	1.6478	--
	(2.2201)	
Sex (female)	-0.8773**	--
	(0.4045)	
Age	-0.0265	--
	(0.0853)	
Dis_Rare	-0.3143	--
	(0.6047)	
Dis_Com	0.5671	--
	(0.4204)	
Health	0.2374	--
	(0.6107)	
Faculty/Staff	0.878	--
	(1.8086)	
Graduate student	0.8862	--
	(1.0383)	
Part-time employed	0.0516	--
	(0.4173)	
Full-time employed	-1.6451	--
	(1.5174)	
Parents own house	0.1165	--
	(0.5448)	
Father unemployed	-2.2895***	--
	(0.8284)	
Mother unemployed	0.6392	--
	(0.4871)	
RESP	1.0317**	--
	(0.5046)	
Average class probabilities	0.458	0.542

(b) Attribute Estimates, by Class	Class1	Class 2
Common disease (intercept)	1.4124***	0.1021
	(0.1972)	(0.0718)
Total Budget (TC)	-0.0004	-0.0071***
	(0.0023)	(0.0007)

Cost-per-patient (CP)	0.0052	-0.0078***
	(0.0037)	(0.0014)
Severity of Disease (SEV)	1.321***	1.1154***
	(0.3208)	(0.1114)
Life-years Gained (LYG)	0.2017***	0.1823***
	(0.0247)	(0.0076)
TC*CP	-0.0001	0.0002***
	(0.0001)	(0.000)
CP*SEV	-0.0035	0.0016
	(0.0033)	(0.0013)
SEV*LYG	-0.0389*	0.0263***
	(0.0228)	(0.01)
TC140	-0.0079	-0.0013
	(0.0061)	(0.0018)
CP80	0.0138	0.0024
	(0.0169)	(0.006)
SevSer	0.1122	-0.191**
	(0.3815)	(0.0954)
LYG5	-0.076	-0.5335***
	(0.1174)	(0.0468)
# of individuals	213	
# of obs.	6816	
Log-L	-1629.680	

Standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table 4. Change in the Probability of Choosing a Drug Associated with a One-Unit Change in an Attribute

	Class1	Class 2
Common disease (intercept)	0.3041	0.0255
Total Budget (TC)	-0.0005	-0.0007
Cost-per-patient (CP)	-0.0004	-0.0012
Severity of Disease (SEV)	0.2185	0.295
Life-years Gained (LYG)	0.0455	0.0487

Note: changes are computed for unitary changes at the mean for TC, CP, LYG and for discrete changes for frequency and SEV, while keeping the other attributes at their sample mean.

Table 5. Marginal rates of substitution in monetary and life years for Class 1

	Marginal WTP in 1000$ of CP	Marginal WTP in millions of $ TC	Marginal WTP in life years forgone [a]
Moderate to Serious	--	--	-4.64
			(1.24)
A life year gained	--	--	--

Standard errors in parentheses computed through the delta method

[a] WTP calculations are computed at the mean of the attributes. i.e. for the figure in the table we would have $\left[\beta_{SEV} + \beta_{CP*SEV} * Mean(CP) + \beta_{\text{SEV*LYG}} * Mean(LYG)\right] / (-\beta_{LYG})$

Table 6. Marginal rates of substitution in monetary and life years for Class 2

	Marginal WTP in 1000$ of CP	Marginal WTP in millions of $ TC	Marginal WTP in life years forgone
Moderate to Serious	174.64	190.48	-7.43
	(33.01)	(21.33)	(0.44)
A life year gained	26.16	27.44	--
	(4.6)	(2.81)	

Standard errors in parentheses computed through the delta method

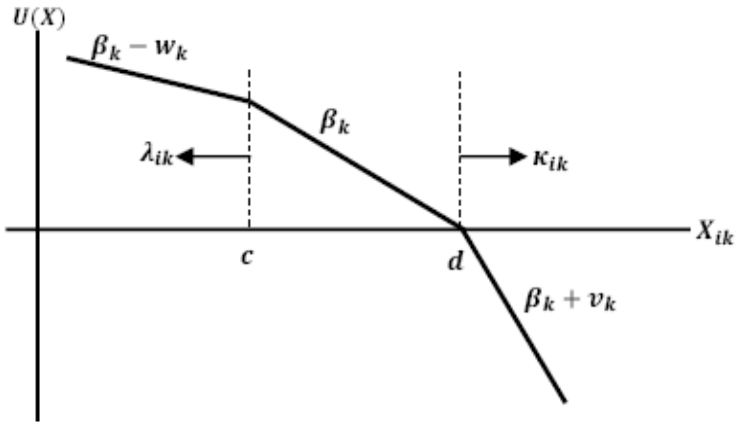

Fig. 1. The impact of cutoff violation on a negatively sloped attribute (e.g. cost) (from Swait, 2001, p. 912).

Appendix 1

Exact description and wording of discrete choice experiment exercise.

In this part of the questionnaire we are interested in your preferences with respect to funding drugs. Each scenario will present information about two health conditions, a drug used to treat each of them, the effect of each drug on a patient's health, and the cost of each drug. For each scenario you will be asked to choose which drug you would prefer that the government include within its public drug plan. If a drug is funded by the public plan patients with the condition can obtain the drug free of charge. The two drugs, the diseases they treat, and the individuals who suffer from the diseases differ only with respect to the attributes listed in the scenario. All other aspects of the decision problem should be assumed identical across the two choice options.

Example Scenario

Imagine that you are a health care decision maker on an Ontario government committee that has been asked to decide which of two drugs will be included within the public drug insurance program. Both drugs are used to treat conditions that arise in the general population. When answering the question below, assume that the characteristics of patients who develop the respective diseases are identical in all respects (e.g., age, marital status, income, education, etc.) except those explicitly mentioned. The two drugs are also identical except with respect to attributes described below; neither drug is associated with adverse side-effects. The money used to fund the chosen drug will come from the provincial public health care budget, which in 2008 was $46 billion. Only one drug can be funded. Please indicate whether you prefer to fund drug A or drug B by placing a tick on one of the boxes below. There are no right or wrong answers.

Example choice problem:

	Drug A used to treat a Common disease	**Drug B used to treat a Rare disease**
The cost of treating a single patient	$10,000 per patient	$12,000 per patient
The total cost of funding the drug program	$20 million to fund	$100 million to fund
The severity of the disease without treatment	Serious Impact	Moderate Impact
The impact of drug treatment on a patient's health	Gain of 1 year	Gain of 10 years

Which drug program would you prefer?	Prefer to fund drug A	Prefer to fund drug B
(*tick one box only*)	☐	☐

Appendix 2

Exact wording of the five characteristics' descriptions.

- The Frequency of the disease in the population

This attribute indicates how often the disease occurs in the population. This can take on two values:

Rare disease: the disease is rare, with less than one case occurring each year for each 2000 people in the population. In a population of 10 million people (the approximate population of Ontario), fewer than 5000 cases would occur each year.

Common disease: the disease is common, with more than one case occurring each year for each 2000 people in the population. In a population of 10 million people (again, the approximate population of Ontario), more than 5,000 cases would occur each year.

- The Cost of treating a single patient

This attribute indicates the total one-time treatment cost for a single patient. All treatment costs arise during a three-month treatment period; no additional treatment costs occur once this is completed. There are seven possible cost values:

- $1,000
- $5,000
- $10,000
- $12,000
- $15,000
- $50,000
- $100,000

- The Total cost of funding the drug program

This attribute indicates the total annual cost to the Ministry of Health if the drug program is funded publicly. There are seven possible costs should a drug be funded by the public plan:

- $5 million total cost
- $10 million total cost
- $20 million total cost
- $50 million total cost
- $100 million total cost
- $150 million total cost
- $200 million total cost

- The severity of the disease without treatment

This attribute indicates the seriousness of a disease if the patient does not obtain treatment. All the diseases you will be asked to consider strike people in their mid-40s. In all cases, once diagnosed, patients who do not receive treatment die prematurely; without treatment, they can expect to live 10 years following diagnosis. Each disease also affects a person's quality of life. The impact on quality of life can take on two values (even after taking into account the effect of miscellaneous services patients seek in an effort to relieve symptoms):

Serious Impact: Patients who suffer from this disease have trouble with their mobility, need help with daily activities and suffer from strong pain. Overall, when asked to rate their health as excellent, good, fair or poor, most patients with the disease rate their health as poor.

Moderate Impact: Patients do not have trouble with their mobility or daily activities, but suffer from minor pain, discomfort, and need rest until their complaints disappear. Overall, when asked to rate their health, most patients with the disease rate their health as good.

Note: Regardless of whether the impact is serious or moderate, patients who are not treated with a drug consume the identical dollar value of miscellaneous health care services in an effort to alleviate their symptoms.

- The Impact of Drug Treatment on a Patient's Health

The drug treatment improves both quality of life and expected length of life. In all cases, the drug returns recipients to excellent health-related quality of life while alive. The impact of the drug treatment on life expectancy takes on four possible values:

- Patients treated with the drug expect to live 15 years longer than those with the disease who do not receive treatment. That is, rather than live only 10 years after diagnosis, a patient treated with the drug can expect to live 25 years.
- Patients treated with the drug can expect to live 10 years longer than those with the disease who do not receive treatment. That is, rather than live only 10 years after diagnosis, a patient treated with the drug can expect to live 20 years.

- Patients treated with the drug can expect to live 5 years longer than those with the disease who do not receive treatment. That is, rather than live only 10 years after diagnosis, a patient treated with the drug can expect to live 15 years.
- Patients treated with the drug can expect to live 1 year longer than those with the disease who do not receive treatment. That is, rather than live only 10 years after diagnosis, a patient treated with the drug can expect to live 11 years.

Appendix 3

In this part of the questionnaire we ask you to indicate your agreement or disagreement with each of a series of statements listed below. Please read each statement carefully before responding, as differences between some statements are small but important. For each statement, please indicate whether you agree or disagree by placing a tick in the appropriate box.

	I Agree	I Disagree
1. The government should not fund drug treatments for conditions with total cost higher than $140 million.	☐	☐
2. The government should not fund drug treatments for conditions with cost per patient higher than $80,000.	☐	☐
3. The government should not choose to fund drug treatments that extend life less than 5 years.	☐	☐
4. The government should only fund drug treatments that have a serious impact on the health status of the patient.	☐	☐