The Patient

Patient-Centered Outcomes Research

Wolters Kluwer Health | Adis

Published in association with:

JOHNS HOPKINS BLOOMBERG SCHOOL *of* PUBLIC HEALTH

Patient 2008; 1 (4): 299-307
1178-1653/08/0004-0299/$48.00/0

Survey-Design and Analytical Strategies for Better Healthcare Stated-Choice Studies

F. Reed Johnson[1] and *Carol Mansfield*[2]

1 Health Solutions, RTI International, Research Triangle Park, North Carolina, USA
2 Health, Social, and Economic Research, RTI International, Research Triangle Park, North Carolina, USA

Abstract

Stated-choice (SC) surveys, such as conjoint analysis, present some interesting problems for researchers that are not addressed in the traditional survey-development literature. While the constraints imposed by preference theory, the experimental design of the choice sets, and the statistical methods used to analyze choice data all pose challenges for researchers new to SC methods, they also direct such researchers towards techniques that are not possible with more traditional survey methods. In this article, we focus on issues of preference heterogeneity (variation in preferences across subjects by observable and non-observable co-variates) and attribute dominance to illustrate the synergistic roles that survey-design and analytical strategies play in SC research. In this article, we show how advanced analytical techniques are likely to be more important than survey-design solutions when addressing preference heterogeneity. Good practice supports the use of mixed-logit and similar modeling approaches to mitigate the problem of unobserved preference or variance heterogeneity. However, if the sample size is not large enough or the survey instrument does not contain questions about important subject characteristics, then the source of heterogeneity cannot be identified and the problems caused by heterogeneity will be magnified.

In contrast, minimizing attribute dominance and testing for attribute dominance relies on careful survey design, rather than more complex analysis. In general, survey design needs careful attention from researchers. No amount of complex analysis can compensate for a poor survey design that can generate only flawed SC data.

If you want to know what people think – why not ask them? It sounds simple, but using surveys to assess patient and provider preferences requires knowledge of both survey design methods and statistical estimation techniques. Health researchers have begun to use stated-choice (SC) survey methods for research on a wide range of issues, including treatment preferences, trade-offs between efficacy and risk, demand for new medications, and the relative importance of treatment characteristics. The growing popularity of SC surveys, particularly discrete-choice experiments (conjoint analysis surveys), can be explained in part because SC surveys offer a systematic method of collecting data that is compatible with economic theory and based on a statistically sound experimental framework.

SC surveys require the same attention to basic survey design and administration procedures as any other survey. However, they also present some interesting problems that are not addressed in the traditional survey-development literature.[1] The constraints imposed by preference theory, the experimental design of the choice sets, and the statistical methods used to analyze choice data pose challenges for inexperienced researchers. However, these same constraints also offer survey-design and analytical solutions to common SC-study problems that are not possible with other survey methods.

Suppose you have completed one of your first SC surveys, which was designed to offer insight into patient treatment preferences or to inform a health-policy decision. You selected the attributes and levels by consulting the clinical literature, you identified the experimental design in a catalog, and you analyzed the data using simple conditional logit. What problems or puzzles did the results present? What should you do differently next time? This is the starting point for our discussion.

While better survey design and analytical strategies may offer separate solutions to many SC problems, researchers who understand preference theory and conjoint analysis methods can combine strategies effectively to take full advantage of the strengths of SC surveys. There are many challenges facing SC researchers. We focus here only on two areas of particular concern for obtaining valid preference results: preference heterogeneity and attribute dominance. We use preference heterogeneity and attribute dominance to illustrate the individual and complementary roles that survey-design and analytical strategies play in SC survey research.

1. Introduction

Our goals for this article are modest. We hope to provide advice and ideas to researchers to deal with two common problems (attribute dominance and preference heterogeneity) and to demonstrate the roles of survey instrument design and analysis techniques. As with any area of research, the more you learn about SC conjoint analysis, the more questions you have. Advanced discussions of experimental design and data estimation require a strong technical background, and we do not attempt to resolve divergent expert opinions that are the focus of active research in the field. However, before we get started, we present a short discussion on two such topics of which SC researchers should be aware – the role of scale versus preference, and experimental design.

1.1 Scale versus Preference Parameters

Choice-model coefficient estimates consist of preference and scale components (equation 1):

$$\hat{\beta} = \mu \times \beta \equiv \frac{\beta}{\sigma} \tag{Eq. 1}$$

where β is a preference parameter and μ is a scale parameter equal to the inverse of the standard deviation of the random-utility error term, σ. Thus, scale indicates the variance in preference estimates that is confounded with the actual preference parameters of interest. Scale can differ between two respondent samples, among respondents, or even among questions or between attributes for the same respondent. In many cases, the data do not provide enough information to separately identify β and μ, so we often assume that $\mu = 1$. However, potential scale differences can cause a number of problems in interpreting and evaluating choice estimates.[2-5]

Researchers must take care not to attribute differences in $\hat{\beta}$ to preference differences when part or all of the difference may simply be a result of differences in scale, especially when comparing across models. Swait and Louviere[4] suggest a strategy for testing for preference differences between two samples with different scales.

A common strategy is to eliminate scale effects by dividing all parameters by one of the estimates, say (the absolute value of) the cost estimate, which converts parameter estimates into monetary units, that is, estimates of marginal willingness to pay (WTP). In equation 2, scale, σ, is eliminated by dividing the preference parameter for individual i, β_i, by the negative of the parameter for the cost estimate, $-\beta_{cost}$. Marginal WTP can then be compared directly between models:

$$\tilde{\beta}_i = \frac{\hat{\beta}_i}{-\hat{\beta}_{cost}} = \frac{\beta_i/\sigma}{-\beta_{cost}/\sigma} = \frac{\beta_i}{-\beta_{cost}} \qquad \text{(Eq. 2)}$$

In theory, any attribute could be selected as the denominator; for example, in Johnson et al.,[6] the probability of a serious adverse event is the denominator resulting in the calculation of maximum acceptable risk, rather than WTP. Of course, this strategy requires assuming $\sigma_i = \sigma$ for all i.

Sorting out the relative role of preference and scale variation can be controversial, and different researchers have approached this problem in different ways.[4,5] [1] Although the focus of this discussion is primarily on strategies to estimate valid preference measures, we will also highlight where scale differences might affect conclusions.

1.2 Experimental Design

The experimental design defines the trade-offs required in SC conjoint surveys and sets them apart from contingent value surveys and other stated-preference or market research techniques. The experimental design allows researchers to obtain valid estimates of individual marginal effects that are not confounded by correlations with other factors in the trade-off tasks. Given the importance of a good experimental design, researchers need to understand some basic principles, such as why you need an experimental design, the criteria used to judge designs, how the number of attributes and levels affects the design, and how the design relates to sample size. Researchers have developed a variety of approaches to experimental design; however, there is still disagreement among experts on the best approach to experimental design.

D-efficiency is the usual standard by which the quality of a design is judged. In theory, maximizing D-efficiency minimizes the average standard error of the estimated coefficients. A higher D-efficiency score will be associated with a design that is more balanced and orthogonal.[1,7-9] Other factors, such as optimal utility imbalance, can also improve experimental design.[10]

A number of programs are available to construct experimental designs. For example, SAS, Sawtooth Software, SPSS, and some other statistical packages provide programs that will create experimental designs. A website at the Sydney University of Technology generates designs based on a cyclical algorithm.[9] Features of the attributes and levels, parameter estimates from previous studies, or prior information about relative importance of attribute levels such as lower or higher costs, restrictions on feasible combinations of attribute levels, and strategies for reducing cognitive burden, such as allowing levels of some attributes to be constant in some choice sets (called overlap), may affect the best approach to experimental design for a given survey.

2. Preference Heterogeneity

2.1 The Problem

Preference heterogeneity among respondents is one of the most interesting features of stated-preference data. Preferences inevitably vary systematically across subjects by observable and non-observable co-variates. Individuals have different preferences among attribute levels and across attributes, and the amount of disagreement across individuals will vary by attribute. As discussed in section 1, there could be heterogeneity across individuals in both preference and scale.

Ex ante knowledge of the possible sources of heterogeneity can inform both the design of the survey and the best approach for data analysis. Simple choice models such as conditional logit obtain preference-parameter or part-worth values for each attribute level in the experimental design. The estimated values (coefficient and scale

1 Figure 1 in the supplementary material (see 'ArticlePlus' at http://thepatient.adisonline.com) illustrates the problem of disentangling preference from scale.

parameter) reflect the mean relative importance of the profile features for the survey sample. The parameters can be biased in the presence of unobserved co-variates, and they do not provide information about the distribution of preferences across the sample.

2.2 How the Problem Could Affect Your Results

In the face of preference heterogeneity, the assumption that all subjects have the same preferences has consequences for both the validity of the parameter estimates and the interpretation of the results for decision making. The level of heterogeneity affects the optimal selection of the attributes and attribute levels in the survey design phase. If two groups of subjects have strong and divergent preferences for an attribute, the mean effect of the attribute may be unrepresentative of either group of subjects. Moreover, a limitation of using a single experimental design that works for the 'average' subject is that the design may not work for subjects in the tails of the preference distribution. For example, the levels of the cost attribute may not be high enough to induce higher-income subjects to pay attention to cost, but may be so high that lower-income subjects focus too much on cost.

Finally, if systematic variation in preferences by characteristics such as age, education, income, or health history are not identified (or are misidentified as mean rather than variance heterogeneity or vice versa), policy-relevant or clinically important nuances may be missed, diminishing the value of the information to decision makers.

Even if we control for observed co-variates, unobserved sources of preference heterogeneity in the sample will bias the parameter estimates. Suppose, for example, that the preferences of patients with diabetes mellitus for glucose control depend on an observed variable, that is, whether they have a close relative with serious diabetes-related sequelae. We can model the effect of that co-variate on choice probabilities. However, suppose that preferences for glucose control also depend on an unobserved variable, that is, patients' tolerance for bearing risk. The influence of that variable will be confounded with the error term, which now is no longer independent and random, but is correlated with the preference parameter for glucose control. The result is the usual bias associated with omitted variables.

2.3 Survey-Design Approaches to Preference Heterogeneity

The most basic protection against unknown or unexpected preference heterogeneity comes in the design phase of the survey. Background research, discussion with experts, and, most importantly, careful pre-testing provide crucial information about subject preferences (see Mansfield and Pattanayak[11] for a discussion of SC survey planning). The design phase for the survey instrument needs to focus on identifying heterogeneity that, if unaccounted for, will bias the estimated coefficients.

In the case of identifiable divergence in sensitivity across respondents to particular attribute levels such as cost, including more levels for the 'cost' attribute that vary over a wider range may help mitigate this problem. With some sacrifice of statistical efficiency, offering more cost levels ensures enough variation at both lower and higher cost levels to induce trade-offs for both lower- and higher-income subjects. Tailored designs that use some information about subjects are also a possibility.

SC surveys on health topics often face the problem that the subjects do not all start at the same baseline, which in turn affects their preferences and risk perceptions. In addition to a standard battery of demographic and economic variables, you may want to include questions on treatment experience, health history, attitudes toward risk, insurance status, and other factors that could help explain preference heterogeneity.

2.4 Analysis Approaches to Preference Heterogeneity

If information on the importance of particular co-variates is known from prior studies, a tempting solution is to estimate split-sample models for each group. If samples are sufficiently large, they can be split into sociodemographic subsamples, estimated separately, and tested for statistical differences. The risk of this strategy is that differences will be attributed to the splitting variable when that variable is correlated with a different causal factor. For example, if education and income are correlated, splitting on education may attribute differences to education, when in fact a model that controls for income may find that education is insignificant.

The specification of utility for diabetes patient k (U_k) is as shown in equation 3:

$$U_k = V_k + \varepsilon_k = \beta_{A1c}A1C + \beta_{hyp}HYP + \beta_c COST + \varepsilon_k \quad \text{(Eq. 3)}$$

where V is the deterministic component of utility, ε is the unobserved component, the βs are the estimated preference parameters, A_{1c} is glucose control, HYP is frequency of hypoglycemic events, and COST is monthly out-of-pocket treatment cost. The basic conditional-logit specification for the probability that subject k will choose alternative i from J alternatives is as shown in equation 4:

$$\mathrm{Prob}(C_k = i) = \frac{\exp(\mu_k V_{ki})}{\sum_{j=1}^{J} \exp(\mu_k V_{kj})} \quad \text{(Eq. 4)}$$

where again, V_{ki} is the deterministic component of utility subject k receives from alternative i, and μ is a scale parameter inversely related to variance, which is often not identifiable and is set equal to 1.

Because individual characteristics are constant within individual choices, variables such as age, income, education, and sex cannot be added linearly to the model specification as they commonly are in regression analysis.

A direct, multivariate approach is to interact individual-characteristic variables with one or more of the A_{1c}, HYP, and COST attributes. Suppose we have information about a subject's close relative that we interact with the glucose-control attribute. The model is then as shown in equation 5:

$$U_k = \beta_{A1c}A1c + \beta_{A1c\times rel}(A1C \times D_{rel}) + \beta_{hyp}HYP + \beta_c COST + \varepsilon_k \quad \text{(Eq. 5)}$$

where D_{rel} is a dummy variable for having a relative with diabetes-related sequelae.

Although interactions with continuous, linear attributes and individual characteristics are straightforward, best practice requires estimating both continuous and categorical attributes as categorical to avoid imposing any functional-form assumptions on utility. Unfortunately, interacting a large number of individual-characteristic variables with multiple effects-coded attributes can result in an intractable number of parameters. Suppose you have five four-level attributes. Omitting one level in each attribute, you need to estimate 15 main-effects parameters. Each interaction with a continuous individual-specific variable adds 15 more parameters to the model. One three-level categorical variable, such as self-reported health status (good, moderate, poor), adds 30 more parameters to the model.

Fortunately, the circumstances of the problem may reduce the dimensionality of the model by indicating where interactions are of primary interest. For example, one might hypothesize that only efficacy preferences are influenced by subjects' current health status, which adds only six parameters to the model. If the primary goal is to estimate WTP, interacting individual characteristics with the cost variable allows the marginal utility of money to vary across subjects, which produces corresponding variation in WTP. Of course, this specification assumes that marginal rates of substitution among non-price attributes are the same for all subjects.

Mixed-logit models can account for unobserved preference heterogeneity, and hierarchical Bayes (HB) models can even estimate a separate parameter vector for each subject in the sample. Mixed logit (sometimes called random-parameters logit) allows preference parameters to be a combination of a population mean β and an individual-specific stochastic component, η_k, which captures any unobserved source of preference variation. Because the most recent versions of many standard statistical packages include mixed-logit procedures, this is becoming the standard analysis for published SC studies.

HB models extend the information provided by η_k to estimate a parameter vector for each individual in the sample.[5,12] In effect, HB models take the sample parameter estimates from mixed logit as priors for an individual-specific Bayesian update based on the information contained in each subject's choice responses. The quality of the update obviously depends on how much information is available from the individual choices. When the experimental design requires multiple blocks or versions of the survey, the particular segment of the design seen by one subject may not support accurate estimates of individual-level preference parameters.

2.5 Summary

In the case of preference heterogeneity, problems in survey design and lack of information on variables associated with unobserved heterogeneity influence the choice of model specification and analytical techniques. If the sample size is not large enough or the survey instrument does not contain questions about important subject characteristics, then the source of heterogeneity cannot be identified and the problems caused by heterogeneity will be magnified.

Even with good survey design, preference heterogeneity needs to be carefully considered. Survey design can only limit the potential for unobserved heterogeneity that can lead to biased preference estimates. Good practice supports the use of mixed logit or variance-heterogeneity models to mitigate the problem of unobserved preference or variance heterogeneity. In cases where policy recommendations depend on the extent of preference heterogeneity or when a study seeks to demonstrate the importance of variation in opinions across a patient population, analytical techniques such as mixed logit are essential. Hensher et al.,[12] Louviere et al.,[13] and Train[5] discuss analytical approaches if one suspects that variance heterogeneity is the issue.

3. Attribute-Dominant Preferences

3.1 The Problem

Choices that satisfy the fundamental properties of utility theory require that subjects be willing to trade less of one attribute in return for more of another attribute. In nearly every study, there are some subjects who always, or nearly always, select the alternative with the best level of one attribute (or on a single attribute level if the levels are not ordered). There are two possible interpretations of attribute-dominant responses. One possibility is that subjects focus on a single attribute to simplify answering the questions. These subjects are inattentive to the trade-off task and provide no meaningful information about their preferences. Unfortunately, it is difficult to discriminate between such inattentive subjects and subjects who simply feel very strongly about a particular attribute and the design space does not contain profiles that are sufficiently attractive to cause them to trade away from the dominant attribute. These subjects provide meaningful information about their preference for the dominant attribute, but no information about their trade-off preferences among the other attributes in the design.[2]

3.2 How the Problem Could Affect Your Results

True or valid attribute-dominant subjects 'stuff the ballot box' in favor of their preferred attribute. Their apparent strong preference for one attribute will bias parameters upward for the better levels and downward for the poorer levels of that attribute. Inattentive subjects who select based on one attribute as a way to finish the survey more quickly bias results with meaningless data.

3.3 Survey-Design Approaches to Attribute Dominance

If there is reason to suspect that some subjects will strongly prefer a particular attribute, we can

2 Differences in the scale parameter across attributes could also produce response patterns that might look like dominant preferences. For example, if respondents are more familiar with one attribute, this could result in a smaller random utility component for that attribute, which would lead to tighter estimates of the coefficient.

get them to reveal trade-offs in other attributes of the design space by overlapping attribute levels in the dominant attribute. Overlap means that the attribute levels are the same for all alternatives in the choice set. Since subjects cannot choose the profile with the better level of the overlapped attribute, they will have to choose based on their willingness to accept trade-offs among other attributes. For example, if you believe that respondents will dominate on the risk of a serious adverse effect, offering some choices where the risk is the same for each alternative forces them to consider the other attributes. At the extreme, the dominant attribute could be dropped and the introductory text to the conjoint questions could specify the level the dominant attribute takes throughout the survey (which could be varied across subsets of respondents to test the impact of the assumption). However, a dominant attribute is, by definition, an attribute on which respondents place great weight, and we therefore recommend pre-testing to explore use of overlap or expanding the ranges of the other attributes if possible.

Choice tasks that are too complicated or not meaningful may result in subjects simplifying the choice task by choosing on the basis of a single attribute. These choices are not valid indications of strength of preference. Good survey design can minimize the number of subjects who employ this decision heuristic. Possible strategies for detecting invalid dominant preferences include the following.

- Employ introductory self-explication questions to determine whether there is any improvement in a less favored attribute that would cause subjects to accept a reduction in the level of the most favored attribute.
- In computer-administered surveys, it may be possible to detect dominant responses and ask follow-up probe questions to elicit the subject's motives for the observed pattern.
- Response patterns where subjects trade-off among attributes at the beginning of the sequence, but are dominant for the remainder of the sequence may reveal additional uninformative responses that might be classified as valid based on an overall count of dominant frequencies.
- Check the amount of time subjects spent on the survey if it is computer administered.
- Ask respondents to select the worst alternative as well as the best.[14]

Table I summarizes non-trading or dominant-preference patterns in patient surveys for Crohn disease[6] and bipolar disorder.[15] The data in the tables indicate the percentages of respondents who always choose the alternative with the highest level for that attribute. A total of 44 respondents (20%) in the Crohn disease survey[6] focused on the symptom severity attribute. Many subjects with bipolar disease focused exclusively on weight gain, but there were some dominant responses for most of the other attributes as well.[15] The difference in dominance patterns between the two surveys demonstrates the variety of situations a researcher can face. The study results will be strengthened or weakened based on the extent to which these patterns can be modeled and explained. Again, careful pre-testing can uncover patterns of dominance and allow the researcher to design additional questions or split-sample tests to model the observed patterns.

Table I. Dominant preferences in the Crohn disease[6] and bipolar disease[15] surveys

Dominant attribute	Percentage of patients
Crohn disease	
Symptom severity	20
Long-term complications	0
Frequency of flare-ups	0
Treatment requires corticosteroids	0
Tuberculosis risk	<1
Progressive multifocal leukoencephalopathy risk	4
Lymphoma risk	<1
Bipolar disease	
Mania severity	4
Mania frequency	3
Depression severity	19
Depression frequency	0
Weight gain	32
Fatigue	11
Cognition	16
Serious adverse effect risk	15

3.4 Analysis Approaches to Dominant Preferences

There are several analytical options for dealing with dominant preferences. First, most researchers simply ignore the problem, in effect assuming that all responses are valid. This strategy risks obtaining biased estimates. Choice estimates will reflect the information in the data. Profiles containing better levels of a dominant attribute will be selected more frequently, thus implying a heavier decision weight for better levels of that attribute. This bias is mitigated to some extent when subjects dominate on multiple attributes. The other simple solution is to drop all subjects with dominant preferences, in effect assuming that all these responses are invalid. A comparison of inclusive and exclusive models will indicate the overall impact of dominant responses in the data.[16]

A better strategy is to model the observed behavior to obtain dominant-corrected parameter estimates. The first step is to construct a dummy variable for each dominant attribute. The dummy can then be interacted with all levels of the attribute for each subject who had dominant preferences for that attribute. The size and significance of the dominant-preference dummies indicate the influence of dominant responses and the remaining parameters are purged of the potential bias. We often find that the dominant dummies are statistically insignificant unless a relatively large number of subjects have dominant responses on one or two attributes.[15] [3]

3.5 Summary

In contrast to the problem of heterogeneous preferences, minimizing attribute dominance and testing for attribute dominance relies on careful survey design, rather than more complex analysis. Survey design has a larger role to play both in creating an instrument that can be used to identify dominant preferences and designing tests for inattentive subjects. Attribute dominance also requires more analyst judgment. Aside from a few obvious cases, such as a respondent who finishes the survey in 30 seconds and selects choice A every time, it may be hard to separate inattentive subjects from those who truly have strong preferences. In addition, respondents may display attribute dominance on some but not all of the choices. Again, the survey-design process provides more information than using different analysis approaches. Careful background research and pre-testing often provides clues as to how subjects are likely to react to the trade-off tasks and may suggest modifications to the survey that will improve data validity. No amount of complex analysis can compensate for a poorly designed instrument that generates fundamentally flawed preference data.

4. Conclusion

Successful preference research is a combination of art and science. Survey design and statistical analysis often provide complementary strategies for dealing with common problems, but some problems are best solved through survey design, while analysis is more important for other problems. We have presented examples of two problems: preference heterogeneity and attribute dominance. More advanced analytical techniques are likely to be more important for preference heterogeneity, while survey-design techniques are essential for minimizing the chance of obtaining flawed data.

Journal reviewers rarely reject manuscripts because of problems with the statistical analysis. Such problems can be fixed with help from colleagues with quantitative skills. Unfortunately, no amount of statistical expertise can compensate for fatally flawed preference data resulting from a poorly designed survey instrument. There is no substitute for observing and listening to actual subjects trying to answer your survey questions during pre-test interviews. They will help you assess how tractable the choice tasks are and help

3 If the dominance pattern results from differences in variance (scale) rather than means (preference parameters), the standard errors on the dummy variables will be incorrect.

you devise ways of reducing the likelihood that a large proportion of your subjects will dominate on a single attribute or otherwise answer the trade-off questions in ways that are not truly informative of their preferences.

Acknowledgments

No sources of funding were used to assist in the preparation of this article. The authors have no conflicts of interest that are directly relevant to the content of this article.

References

1. Kanninen B, editor. Valuing environmental amenities using choice experiments: a common sense guide to theory and practice. Dordrecht: Springer, 2007
2. Lancsar E, Louviere J, Flynn T. Several methods to investigate relative attribute impact in stated preference experiments. Soc Sci Med 2007; 64 (8): 1738-53
3. Louviere J, Street D, Carson R, et al. Dissecting the random component of utility. Market Lett 2002; 13 (3): 177-93
4. Swait J, Louviere J. The role of the scale parameter in the estimation and comparison of multinomial logit models. J Market Res 1993; 30 (3): 305-14
5. Train K. Discrete choice methods with simulation. Cambridge (UK): Cambridge University Press, 2003
6. Johnson FR, Özdemir S, Mansfield C, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. Gastroenterology 2007; 133 (3): 769-79
7. Chrzan K, Orme B. An overview and comparison of design-based strategies for choice-based conjoint analysis. Sequim (WA): Sawtooth Software Research Paper Series, 2000 [online]. Available from URL: http://www.sawtoothsoftware.com/download/techpap/desgncbc.pdf [Accessed 2008 Oct 30]
8. Kuhfeld WF. Experimental design, efficiency, coding and choice design. SAS Technical Support TS-722C, 2005 [online]. Available from URL: http://support.sas.com/techsup/technote/ts722.pdf [Accessed 2008 Nov 6]
9. Centre for the Study of Choice (CenSoC). Discrete choice experiments [online]. Available from URL: http://crsu.science.uts.edu.au/choice [Accessed 2008 Nov 10]
10. Kanninen B. Optimal design for multinomial choice experiments. J Market Res 2002; 39 (2): 214-27
11. Mansfield CA, Pattanayak SK. Getting started. In: Kanninen B, editor. Valuing environmental amenities using choice experiments: a common sense guide to theory and practice. Dordrecht: Springer, 2007
12. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. Cambridge (UK): Cambridge University Press, 2005
13. Louviere J, Swait J, Hensher D. Stated choice methods: analysis and application. Cambridge (UK): Cambridge University Press, 2000
14. Marley AAJ, Louviere JJ. Some probabilistic models of best, worst, and best-worst choices. J Math Psychol 2005; 49: 464-80
15. Johnson FR, Ozdemir S, Manjunath R, et al. Factors that affect adherence to bipolar disorder treatments: a stated-preference approach. Med Care 2007; 45 (6): 545-52
16. McIntosh E, Ryan M. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: implications of discontinuous preferences. J Econ Psychol 2002; 23: 367-82

Correspondence: Dr *Carol Mansfield*, Health, Social, and Economic Research, RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709, USA.
E-mail: carolm@rti.org

