

available at www.sciencedirect.com

SCIENCE @ DIRECT®

journal homepage: www.elsevier.com/locate/jval

SCIENTIFIC REPORT

Conjoint Analysis Applications in Health—a Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force

John F. P. Bridges, PhD^{1,*}, A. Brett Hauber, PhD², Deborah Marshall, PhD³, Andrew Lloyd, DPhil⁴, Lisa A. Prosser, PhD⁵, Dean A. Regier, PhD⁶, F. Reed Johnson, PhD², Josephine Mauskopf, PhD⁷

¹Department of Health Policy & Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ²Health Preference Assessment Group, RTI Health Solutions, Research Triangle Park, NC, USA; ³Department of Community Health Sciences Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada; ⁴Patient Reported Outcomes Group, Oxford Outcomes, Oxford, UK; ⁵Child Health Evaluation and Research Unit, Department of Pediatrics, University of Michigan, Ann Arbor, MI, USA; ⁶Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; ⁷Health Economics Group, RTI Health Solutions, Research Triangle Park, NC, USA

ABSTRACT

Background: The application of conjoint analysis (including discrete-choice experiments and other multiattribute stated-preference methods) in health has increased rapidly over the past decade. A wider acceptance of these methods is limited by an absence of consensus-based methodological standards. **Objective:** The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Good Research Practices for Conjoint Analysis Task Force was established to identify good research practices for conjoint-analysis applications in health. **Methods:** The task force met regularly to identify the important steps in a conjoint analysis, to discuss good research practices for conjoint analysis, and to develop and refine the key criteria for identifying good research practices. ISPOR members contributed to this process through an extensive consultation process. A final consensus meeting was held to revise the article using these comments, and those of a number of international reviewers. **Results:** Task

force findings are presented as a 10-item checklist covering: 1) research question; 2) attributes and levels; 3) construction of tasks; 4) experimental design; 5) preference elicitation; 6) instrument design; 7) data-collection plan; 8) statistical analyses; 9) results and conclusions; and 10) study presentation. A primary question relating to each of the 10 items is posed, and three sub-questions examine finer issues within items. **Conclusions:** Although the checklist should not be interpreted as endorsing any specific methodological approach to conjoint analysis, it can facilitate future training activities and discussions of good research practices for the application of conjoint-analysis methods in health care studies.

Keywords: conjoint analysis, discrete-choice experimental, economic evaluation, good research practice.

Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Background to the task force report

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Preference-based Methods Special Interest Group's Conjoint Analysis Working Group developed a proposal for a task force on conjoint-analysis good research practices. With the increase in use of conjoint analysis, a structure to guide the development, analysis, and publication of conjoint analyses in health care studies would be useful for researchers, reviewers, and students. The task force proposal was submitted to the ISPOR Health Science Policy Council (HSPC) in November 2008. The HSPC recommended the proposal to the ISPOR Board of Directors where it was subsequently approved in January 2009.

The ISPOR Conjoint Analysis Good Research Practices Task Force met regularly via teleconference and in person at ISPOR meetings to identify the important steps in a conjoint analysis,

to develop and refine the key criteria for good research practices, write the outline, and draft the subsequent report.

ISPOR members and invited international experts contributed to the consensus development of the task force report via comments made during a Forum presentation at the 2009 ISPOR 14th Annual International Meeting (Orlando, FL, USA), through the comments received from the draft report's circulation to the Conjoint Analysis Reviewer Group and an international group of reviewers selected by the task force chair.

The task force met in person in September 2009 to discuss and come to consensus on the more controversial issues that arose. The draft report was revised as appropriate to address comments from these review opportunities. The final step in the consensus process was circulation of the report to the ISPOR membership in September 2010 with an invitation to review and comment. A total of 42 reviewers submitted written or verbal comments from all occasions to review the draft report.

* Address correspondence to: John F. P. Bridges, Assistant Professor, Johns Hopkins Bloomberg School of Public Health, 624 North Broadway, Rm 689, Baltimore, MD 21205, USA.

E-mail: jbridges@jhsph.edu

1098-3015/\$36.00 – see front matter Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

doi:10.1016/j.jval.2010.11.013

In October 2010, the task force met in person one last time to finalize the report following the ISPOR membership review. All comments are posted on the task force's webpage, and reviewers acknowledged. To view comments, background and membership of the task force, please visit the Conjoint Analysis in Health Care webpage: <http://www.ispor.org/TaskForces/ConjointAnalysisGRP.asp>.

Introduction

Understanding how patients and other stakeholders value various aspects of an intervention in health care is vital to both the design and evaluation of programs. Incorporating these values in decision making may ultimately result in clinical, licensing, reimbursement, and policy decisions that better reflect the preferences of stakeholders, especially patients. Aligning health care policy with patient preferences could improve the effectiveness of health care interventions by improving adoption of, satisfaction with, and adherence to clinical treatments or public health programs [1–3].

Economists differentiate between two approaches to the measurement of preferences: revealed and stated [4]. Both approaches stem from the same theoretical foundation. Revealed preferences are, however, derived from actual observed market activities and require researchers to use complicated econometric methods to identify them. Stated preferences are derived from surveys and allow researchers to control the way in which preferences are elicited.

In health, the term “preferences” includes methods beyond the stated and revealed preference paradigms. For example, methods such as the time-trade-off or standard gamble, which are used to calculate quality-adjusted life years (QALYs), are referred to as preference-based. Such methods are based on cardinal utility and are beyond the scope of a scientific report on stated preferences. Stated-preference studies are preferable to QALY or attitudinal-based valuation methods because they are grounded in consumer theory and the psychology of choice.

Stated-preference methods fall into two broad categories:

- Methods using ranking, rating, or choice designs (either individually or in combination) to quantify preferences for various attributes of an intervention (often referred to as conjoint analysis, discrete-choice experiments, or stated-choice methods), or
- Methods using direct elicitation of monetary values of an intervention (including contingent valuation or willingness-to-pay and willingness-to-accept methods) [4,5].

A simple distinction between these two categories is that the latter aims to estimate demand for a single product, whereas the former aims to explore trade-offs between a product's attributes and its effect on choice. In practice, the distinctions between the two categories have blurred, with researchers estimating demand using multiple-question and discrete-choice formats, and researchers using preference estimates to calculate willingness-to-pay for attributes.

This scientific report focuses on the first of these approaches. Following standard convention in health care, we refer to them as conjoint analysis. However, we acknowledge that many others would prefer the term “discrete-choice experiment” over “conjoint analysis.” This said, most of the material in this report applies equally to discrete-choice experiments and other types of conjoint analysis.

Conjoint analysis in health care studies

There has been a rapid increase in the application of conjoint analysis in health care studies [6–8]. Conjoint analysis is a decomposition method, in that the implicit values for an attribute of an intervention are derived from some overall score for a profile consisting (conjointly) of two or more attributes [9–13].

Conjoint-analysis methods are particularly useful for quantifying preferences for nonmarket goods and services or where market choices are severely constrained by regulatory and institutional factors, such as in health care [14]. Conjoint analysis has been applied successfully to measuring preferences for a diverse range of health applications. Examples include cancer treatments [15,16]; human immunodeficiency virus prevention [17], testing [18], and treatment [19]; dermatology services [20]; asthma medications [21]; genetic counseling [22,23]; weight-loss programs [24]; diabetes treatment [25] and prevention [26]; colorectal cancer screening [3,8]; depression [27]; and treatments for Alzheimer's disease [28].

The potential benefits of conjoint analysis go beyond the valuation of health care interventions. Increasingly, conjoint analysis also is used as a means to understand patient preferences for health states and as a means to value the various health states described by patient-reported outcomes and health-related quality-of-life scales [29,30]. Licensing authorities recently have taken an interest in conjoint analysis to assess patients' willingness to accept the therapeutic risks associated with more effective new treatments [31]. Conjoint analysis also offers a mechanism for patients to participate in decision making [32,33] and may facilitate shared decision making [34]. Conjoint analysis also can be used to understand clinical decision making [35] and how different stakeholders may value outcomes [36].

The task force thus endeavored to provide broad guidance on good research practices by suggesting a structure to guide the development, analysis, and publication of conjoint analyses in health care studies, without necessarily endorsing any one approach. For its report, the task force also decided to use the checklist format to guide research [37], rather than state the principles by which such research should be conducted [38,39].

How the checklist can be used

The checklist should be used to understand the steps involved in producing good conjoint-analysis research in health care. The final format of the checklist follows the format established by Drummond and colleagues [40]. By outlining a systematic process of good research practices for applying conjoint analysis—from formulating the research question through the presentation of the results (either in presentations, abstracts, reports, or manuscripts)—we intend to facilitate the research process and to highlight important issues that often are neglected or poorly executed. We highlight “good research practices” rather than “best research practices,” with many elements of the checklist presented as methodological considerations rather than as a necessary or sufficient argument for research excellence.

Description of the checklist

The findings of the task force are presented as a 10-item checklist and summarized in Figure 1. The checklist includes items relating to the: 1) research question; 2) attributes and levels; 3) construction of tasks; 4) experimental design; 5) preference elicitation; 6) instrument design; 7) data-collection plan; 8) statistical analyses; 9) results and conclusions; and 10) study presentation.

Implicit in the structure of the checklist is that some tasks should be considered jointly or collectively. These joint tasks

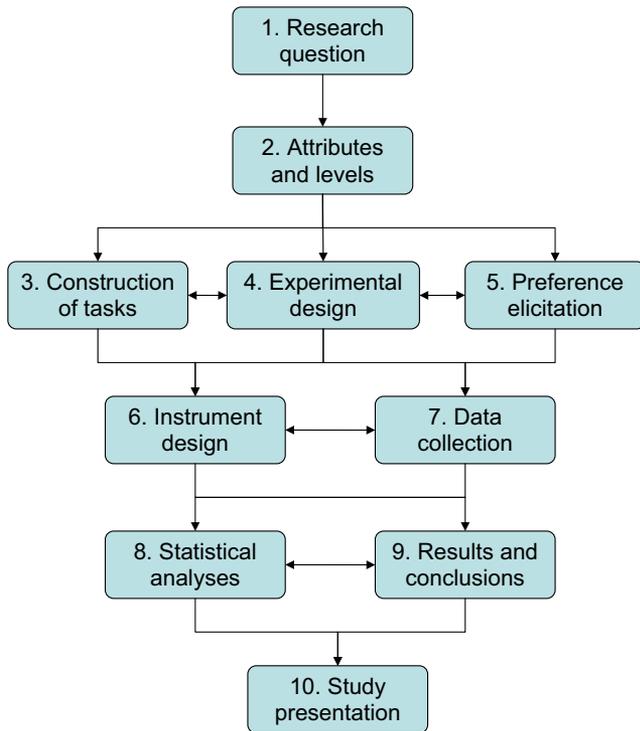


Fig. 1 – A checklist for conjoint analysis in health care.

are arranged horizontally in Figure 1. For example, in constructing the preference-elicitation tasks, experimental design and preference-elicitation methods should be considered as inter-related. Likewise, instrument design is closely related to data collection, and choice of statistical analyses and the ability to draw results and conclusions also are inseparable. More experienced researchers may see additional connections (or may suggest that all 10 items are linked). We highlight these particular relationships to emphasize that the checklist should not be used as a simple “cookbook.”

In the remaining sections of this report, we describe issues to be considered in evaluating each of these 10 items and elaborate on additional points in each section of the checklist. These items are summarized in Table 1. We have kept cross-referencing to a minimum and avoided citing complex articles or books from other disciplines. Thus, we caution readers not to consider this report as an exhaustive reference but simply as an introduction to conjoint-analysis good research practices.

Research question

Following generally accepted research practices in health care, a conjoint-analysis study must clearly state a well-defined research question that delineates what the study will attempt to measure [5]. For example, a conjoint analysis might be undertaken to quantify patients' relative preferences for cost, risk of complications, and health care service location for a given medical intervention. Specifying a testable hypothesis, defining a study perspective, and providing a rationale for the study are important good research practices for applications of conjoint analysis in health care.

Testable hypothesis

In addition to defining the research question, researchers should state any hypotheses to be tested in the study or acknowledge that the study is exploratory and/or descriptive. A testable hypothesis may be implicit in the research question itself. For example, if the

research question is to determine whether changes in surgical wait time influence patient treatment choice, the testable null hypothesis is that the parameter estimate for the wait-time attribute is not statistically significantly different from zero. In other words, the hypothesis test is designed to infer whether a change in the level of the attribute (e.g., a change in surgical wait time from 1 to 2 months) is statistically significant. If the null hypothesis is rejected for a given attribute, then the parameter estimate on that attribute is statistically significant, indicating that it has played a role in the patients' responses.

Study perspective

Researchers should define the study perspective, including any relevant decision-making or policy context. The research question “What are patients willing to pay for treatment to reduce the rate of relapse in multiple sclerosis?” includes both the items to be measured—the trade-off between cost and reduction in relapse rate—and the perspective and decision context of the analysis, i.e., the patient's perspective in making treatment decisions. Here researchers may want to provide even more specifics in defining the study perspective by focusing on a particular type of patient or a particular timing or environment. Although it is good research practice to offer the most accurate study perspective possible, the more specific the perspective, the more difficult it may be to find respondents.

Rational for using conjoint analysis

A conjoint-analysis study should explain why conjoint methods are appropriate to answer the research question. Conjoint analysis is well suited to evaluate decision makers' willingness to trade off attributes of multi-attribute services or products. The multiple sclerosis research question posed in the previous paragraph involves explicit trade-offs between measurable attributes, so it can be answered using conjoint analysis. The research question also could be addressed using alternative methods such as contingent valuation. Researchers should identify not only whether conjoint analysis can be used to answer the research question but also why conjoint analysis is preferable to alternative methods.

Attributes and levels

The objective of conjoint analysis is to elicit preferences or values over the range of attributes and levels that define profiles in the conjoint-analysis tasks. Although all attributes that potentially characterize the alternatives should be considered, some may be excluded to ensure that the profiles are plausible to subjects. For the chosen attributes, the attribute levels should encompass the range that may be salient to subjects, even if those levels are hypothetical or not feasible given current technology. Again, the choice of attribute levels may need to be restricted. Authors should explain both inclusions and omissions of attributes and levels. Good research practices should include attribute identification, attribute selection, and level selection.

Attribute identification

Identifying attributes should be supported by evidence on the potential range of preferences and values that people may hold. Here researchers need to strike a balance between what may be important to the respondent and what is relevant to the particular policy- or decision-making environment. The eventual balance of these competing objectives must be guided by the research question and the study perspective. Sources of evidence to support the inclusion or exclusion of attributes should include literature reviews and other evidence on the impact of the disease and the nature of the health technology being assessed. Consultation with clinical experts, qualitative research [41], or other preliminary studies [42] can provide the basis for identifying the full set of

Table 1 – A checklist for conjoint analysis applications in health care.

1. Was a well-defined research question stated and is conjoint analysis an appropriate method for answering it?
 - 1.1 Were a well-defined research question and a testable hypothesis articulated?
 - 1.2 Was the study perspective described, and was the study placed in a particular decision-making or policy context?
 - 1.3 What is the rationale for using conjoint analysis to answer the research question?
2. Was the choice of attributes and levels supported by evidence?
 - 2.1 Was attribute identification supported by evidence (literature reviews, focus groups, or other scientific methods)?
 - 2.2 Was attribute selection justified and consistent with theory?
 - 2.3 Was level selection for each attribute justified by the evidence and consistent with the study perspective and hypothesis?
3. Was the construction of tasks appropriate?
 - 3.1 Was the number of attributes in each conjoint task justified (that is, full or partial profile)?
 - 3.2 Was the number of profiles in each conjoint task justified?
 - 3.3 Was (should) an opt-out or a status-quo alternative (be) included?
4. Was the choice of experimental design justified and evaluated?
 - 4.1 Was the choice of experimental design justified? Were alternative experimental designs considered?
 - 4.2 Were the properties of the experimental design evaluated?
 - 4.3 Was the number of conjoint tasks included in the data-collection instrument appropriate?
5. Were preferences elicited appropriately, given the research question?
 - 5.1 Was there sufficient motivation and explanation of conjoint tasks?
 - 5.2 Was an appropriate elicitation format (that is, rating, ranking, or choice) used? Did (should) the elicitation format allow for indifference?
 - 5.3 In addition to preference elicitation, did the conjoint tasks include other qualifying questions (for example, strength of preference, confidence in response, and other methods)?
6. Was the data collection instrument designed appropriately?
 - 6.1 Was appropriate respondent information collected (such as sociodemographic, attitudinal, health history or status, and treatment experience)?
 - 6.2 Were the attributes and levels defined, and was any contextual information provided?
 - 6.3 Was the level of burden of the data-collection instrument appropriate? Were respondents encouraged and motivated?
7. Was the data-collection plan appropriate?
 - 7.1 Was the sampling strategy justified (for example, sample size, stratification, and recruitment)?
 - 7.2 Was the mode of administration justified and appropriate (for example, face-to-face, pen-and-paper, web-based)?
 - 7.3 Were ethical considerations addressed (for example, recruitment, information and/or consent, compensation)?
8. Were statistical analyses and model estimations appropriate?
 - 8.1 Were respondent characteristics examined and tested?
 - 8.2 Was the quality of the responses examined (for example, rationality, validity, reliability)?
 - 8.3 Was model estimation conducted appropriately? Were issues of clustering and subgroups handled appropriately?
9. Were the results and conclusions valid?
 - 9.1 Did study results reflect testable hypotheses and account for statistical uncertainty?
 - 9.2 Were study conclusions supported by the evidence and compared with existing findings in the literature?
 - 9.3 Were study limitations and generalizability adequately discussed?
10. Was the study presentation clear, concise, and complete?
 - 10.1 Was study importance and research context adequately motivated?
 - 10.2 Were the study data-collection instrument and methods described?
 - 10.3 Were the study implications clearly stated and understandable to a wide audience?

attributes (and even possible attribute levels) that characterize the profiles to be evaluated.

Attribute selection

The subset of all possible attributes that should be included in the conjoint-analysis tasks can be determined on the basis of three criteria: relevance to the research question, relevance to the decision context, and whether attributes are related to one another. Attributes central to the research question or to the decision context must either be included or held constant across all profiles. It is important to control for any potential attributes that are omitted from the conjoint-analysis tasks but that correlate with attributes that are included in these tasks. In the United States health care market, insurance coverage and out-of-pocket medical expenses for procedures are routine for many patients. Cost may be perceived as correlated with improvements in medical outcomes or with access to advanced interventions. If cost is not included, it should be controlled for by informing subjects that it is constant across profiles.

Discussion with experts and further pilot testing with subjects can be used to narrow the list of attributes. If the number of possible attributes exceeds what one may find possible to pilot in a conjoint analysis, it may prove beneficial to use other types of

rating and/or ranking exercises (often referred to as compositional approaches) to assess the importance of attributes and to facilitate the construction of the final list of attributes to be included.

Level selection

Once the attributes have been decided upon, researchers must identify the levels that will be included in the profiles in the conjoint-analysis tasks. Levels can be categorical (e.g., a public or private hospital), continuous (a copayment of \$10, \$20, or \$30), or a probability (a chance of rehospitalization of 2%, 5%, or 10%). Although an emerging literature discusses the subjective recoding of the levels [43], no clear best practice has emerged to avoid recoding of the levels. Thus, researchers should avoid the use of ranges to define attributes (such as a copayment from \$5–\$10) because this requires the respondent to subjectively interpret the levels, and the resulting ambiguity will affect the results.

Researchers also are cautioned against choosing too many attribute levels. Although some attributes may require more or fewer levels (especially those that are categorical), it is good research practice to limit levels to three or four per attribute. Finally, researchers should avoid the use of extreme values that may cause a grounding effect. Unless it is required for the research question, researchers

need not span the full breadth of possible levels. For example, if one were constructing an attribute to define the distance to the nearest service in a national study, it might be plausible to have very small and very large distances, but these could be considered outliers. Instead one might form levels across the interquartile range or at plus and minus one standard deviation from the mean. Whatever the logic used to determine the levels of an attribute, researchers should make their decision making transparent, and assumptions need to be tested during the pilot testing.

Construction of tasks

Conjoint-analysis tasks are the mechanism by which possible profiles are presented to respondents for the purpose of preference elicitation. Conjoint-analysis tasks can be assembled in a number of ways by varying the numbers of attributes, profiles (options or choices), and other alternatives. Thus, researchers should consider the use of full or partial profiles, an assessment of the appropriate number of profiles per task, and the inclusion of opt-out or status-quo options.

Full or partial profiles

Within the tasks that respondents will evaluate, profiles (alternatives or choices) can be presented with all the attributes that are being considered in the study (a full profile) or with only a subset of the attributes (a partial profile). Although it is generally considered good practice in health care research to work with full profiles, researchers should determine, through qualitative research or pilot testing, whether subjects can reasonably evaluate the full profiles.

If researchers believe that the complexity of the conjoint-analysis task will encourage respondents to develop or employ simplifying heuristics, such as focusing on only a few attributes while ignoring others, then partial profiles may be preferred. If the use of partial profiles is undesirable, then tasks can show full profiles, but researchers should constrain some attribute levels to be the same (i.e., overlap) between the profiles [44].

Number of profiles

Increasing the number of profiles included in each conjoint-analysis task is considered an efficient way to present more profiles to each respondent. However, little has been written in health care research on the effect that increasing the number of profiles has on respondents. The optimal number of tasks also depends on the method of preference elicitation, in addition to the number and complexity of the attributes included in each task. Furthermore, the number of profiles in each task will have implications for the experimental design (see Fig. 1, checklist item 4).

In some studies, subjects may be presented with a set of many alternative profiles and asked to order or rank the profiles from most preferred to least preferred. In this type of study, subjects often complete only one task. In other studies, profiles are grouped into sets, and respondents are asked to choose among the alternatives in each set. In the latter approach, respondents are typically asked to complete multiple tasks and thus evaluate multiple sets. In health care applications, it is common to present only two profiles in each task, often using the forced choice-elicitation format.

Opt-out or status-quo options

In designing conjoint-analysis tasks, researchers may want to incorporate opt-out or status-quo options. An opt-out option allows the respondent to not choose any of the alternatives in the choice set (in health care, this is like choosing a no-treatment option). A status-quo option is comparable to allowing a respondent to choose to keep his or her current treatment. Such options differ from “I can’t choose” or “I am indifferent between the options”—

which generally are considered less desirable or even poor research practices—in that opting out or choosing the status quo involves the elicitation of (strict) preferences.

The inclusion of an opt-out or status-quo option may be inappropriate for many types of research questions in health care. Including these options, however, can be useful, or even necessary, if researchers are assessing the potential demand or market share of a (novel) product. Finally, the inclusion of an opt-out or status-quo option may have serious implications for the experimental design. It will limit the ability of the researchers to estimate the underlying preference structure because the option results in the censoring of data.

Experimental design

A major advantage of conjoint analysis is that it gives researchers control over the experimental stimuli used to generate the preference data. Researchers thus avoid problems of confounding, correlation, insufficient variation, and unobserved variables common in the analysis of revealed-preference data. The experimental design defines the experimental stimuli used to elicit choices or judgments necessary to identify underlying preference relations. Good research practice requires a detailed explanation and justification of the chosen experimental design, an analysis of the properties of the experimental design, and justification for the number of conjoint tasks included in the data-collection instrument.

Choice of experimental design

The goal of a conjoint-analysis experimental design is to create a set of tasks that will yield as much statistical information as possible for estimating unbiased, precise preference parameters (usually preference weights for all attribute levels) [10]. Good designs have several desirable properties. A design is orthogonal if all attribute levels vary independently, and thus are not correlated. A design is *balanced* when each level of an attribute occurs the same number of times. A design is efficient when it has the smallest variance matrix. Efficient designs are orthogonal and balanced if the underlying statistical model assumes linearity. For nonlinear statistical models (e.g., the multinomial logit model), orthogonality and level balance may not result in the most efficient design. Design algorithms seeking to maximize efficiency can be used. D-efficient designs assume a particular form of the variance matrix and seek to minimize average parameter-estimate variances.

Theoretically efficient designs may have undesirable empirical properties. For example, the variance of the design depends on the actual parameter values. Most design programs assume equal preference weights in calculating efficiency. Even when the actual preference weights are unknown, attributes often have a natural ordering such as “mild,” “moderate,” and “severe.” Designs that do not incorporate such information can result in dominated pairs, where all the attribute levels of one alternative are better than the attribute levels of another alternative. Such choices provide no preference information, although they may be included in a theoretically efficient design. Another concern is that designs can include implausible attribute combinations. For example, an orthogonal design might include a profile that combines “severe nausea” with “no restrictions on activities of daily living.” Eliminating the implausible combinations will result in a design that is no longer orthogonal.

The problem of including illogical combinations is a special case of the larger challenge of balancing the goals of minimizing statistical error and minimizing measurement error (see Fig. 1, checklist item 6). Conjoint questions using orthogonal designs may be too difficult or confusing for respondents to answer, so gains from reducing statistical error by using an orthogonal design

may be outweighed by losses from increasing measurement error by increasing the cognitive burden of the task [44].

Researchers have offered several alternative approaches for conjoint studies, including D-optimal and near-D-optimal designs [45,46], utility-imbalanced designs [47], cyclical designs [48], and random designs [49]. Each of these approaches has advantages and disadvantages, including conformity with theoretical optimality, flexibility in accommodating prior information on preferences and constraints on plausible combinations, and ease of construction [13,50].

Properties of the experimental design

There remains no gold standard for experimental design and it is important for researchers to describe, evaluate, and document how the particular design meets the goals of the study. Potential criteria for evaluating designs include the following:

- Efficiency score,
- Correlations among attribute levels,
- Correlations among attribute-level differences,
- Level balance,
- Number of overlapping attributes,
- Restrictions on implausible combinations,
- Cognitive difficulty.

Experimental-design programs such as SAS (Cary, NC), SPSS (Chicago, IL), and Sawtooth Software (Sequim, WA) typically generate a number of design diagnostics to assist in evaluating designs, and at least one website offers a service to measure the relative efficiency of any design [51].

Number of conjoint tasks

Some experimental designs will automatically set the number of conjoint-analysis tasks. However, many computer-generated designs require the researcher to select the number of conjoint-analysis tasks to be generated. In determining the number of conjoint-analysis tasks, three broad questions need to be answered:

- How many tasks will be generated as part of the experimental design?
- What is the maximum number of tasks that a respondent can answer?
- Will the respondent receive any tasks that are not explicitly part of the design?

For large designs it is often necessary to use a “block” design, which partitions the main experimental design into a fixed number of groups. There is still much debate in health care research as to the appropriate number of conjoint-analysis tasks a respondent can complete, but it is good practice to include 8 to 16 conjoint-analysis tasks. Still, some conjoint analysis practitioners advocate that respondents can complete up to 32 tasks.

Preference elicitation

The purpose of conjoint analysis is to measure respondents’ preferences or values within the bounds of specific research questions, hypotheses (if any), and a stated study perspective. Multiple preference-elicitation methods exist, and the appropriate choice needs to be well justified by researchers. Good research practices for the choice of the preference-elicitation method must cover the motivation and explanation of tasks, elicitation format, and any qualifying questions.

Motivation and explanation

It is a good research practice to offer respondents sufficient motivation and explanation to encourage them to respond to all conjoint-analysis tasks; however, very little has been written on the

level of information that is appropriate for applications in health care. Depending on the research question, study perspective, and study population, care must be taken to offer sufficient information to the respondents as to why you want them to complete the tasks. Too much information can result in yeaying, strategic voting, or information overload.

It is good research practice to introduce attributes and levels prior to the introduction of the actual tasks. The respondents’ comprehension can be facilitated or tested with a simple question relating to each attribute as it is introduced. Likewise, the conjoint-analysis tasks can be introduced with an example question that can be complete or which the respondent can fill out as a practice question.

If the research question or study perspective requires the respondent to think of a special circumstance, act as a proxy for another decision maker, or consider an action in the future, then this must be explained clearly and facilitated throughout all of the tasks. Some researchers in health care use “cheap talk,” a motivating statement given during the tasks to ensure that respondents stay focused. The impact of such methods is not well understood in health care applications.

Elicitation format

Multiple question formats can be used in conjoint-analysis studies. Researchers should ensure that the elicitation format used is appropriate to answer the study’s research questions. In addition, data generated using different question formats will require different methods of statistical analysis. In a discrete-choice or forced-choice format, each task includes two or more profiles from which respondents are asked to choose. In recent years, there has been interest in Best-Worst Scaling, with three distinct elicitation formats highlighted in the literature [52]. Researchers should clearly explain why a specific elicitation format was chosen over alternative approaches.

Qualifying questions

The primary purpose of a conjoint-analysis task is for the respondent to state their preference by rating, ranking or choosing among the profiles. Increasingly, researchers are considering the use of other qualifying questions following the preference elicitation. These can ask the respondent to discuss their level of confidence in their answer, estimate their willingness to pay for their chosen profile, estimate their willingness to accept the less preferred option, or compare their chosen outcome with a choosing nothing or status-quo option.

When an opt-out alternative is included in the choice task, some subjects may choose the opt-out option as a means to avoid evaluating the hypothetical alternatives. An alternative to an opt-out alternative in each conjoint-analysis task is to include an opt-out option as a separate question following each task. That is, respondents who select Alternative A in a forced-choice question are then offered Alternative A or the opt-out option in a follow-up question [3].

Instrument design

Conjoint-analysis tasks represent the method of preference elicitation. However, these tasks need to be delivered as part of a larger survey instrument. Good research practices for conjoint-analysis instrument design require researchers to consider the collection of respondent information, the provision of contextual information, and the level of burden. Supplemental to these concerns, researchers should consider more general issues associated with the survey design [53,54].

Respondent information

Respondents' previous knowledge and experience with health outcomes or services may influence their preferences. It is important to elicit respondent-specific health and sociodemographic information to allow for testing for systematic differences in preferences based on these characteristics (e.g., attitudinal, health history and/or status, treatment experience). Respondents' health status also may influence their preferences in a systematic way. This may reduce the generalizability of the findings if not considered as part of the study design [55].

If respondents' preferences vary according to specific characteristics or experiences, identifying these subgroups could be valuable in tailoring programs to specific types of patients or targeting interventions to individual preferences for health outcomes.

Contextual information

The introductory section for the data-collection instrument can present the overall context of the study, describe the attributes and levels that will be included in the conjoint-analysis tasks, and include one or more practice versions of the tasks. It is important to describe all attributes and levels thoroughly and consistently, to ensure that all respondents are evaluating the same task and not making unobservable assumptions about the attributes and levels in a given profile. For example, respondents may have quite different outcomes in mind for symptom levels described simply as mild, moderate, or severe.

Bias may be introduced by the order in which attributes are presented or by the order of the questions. The number of attributes and levels may induce measurement error as well [56]. Work by Kjaer and colleagues [57] suggested that respondents can show differential sensitivity to price, depending on where the cost attribute occurs in the profile. Varying the order of attributes may be prudent. Randomizing the order of tasks is good practice.

Level of burden

If the data-collection instrument is too cognitively burdensome or too long, the likely result is high rates of nonresponse or partial response. Detailed pretests with potential respondents can identify these potential issues early in the development process. Appropriate length for a particular survey will depend on the mode of administration (mail, Internet, etc.), as well as the level of difficulty of the choice tasks. The level of incentive for participation in the survey will vary with the complexity of the survey, survey length, and survey population. It can be tailored to match the context of the survey. It is good practice to provide a respondent with incentive for participation in the survey in a manner that is in compliance with ethical guidelines.

It is important to include face-to-face pretest interviews and a quantitative pilot test as part of the construction of the data-collection instrument. Careful pretesting can identify areas of misunderstanding or common errors, as well as whether the survey is too long. In addition, it can reveal whether respondents understand the instructions and feel the questions are appropriate. A formal pilot study, in which the final questionnaire is administered on a subset of the final sample, allows for consistency or rationality tests and can provide estimates of coefficient size and direction.

Data collection

Given that conjoint analysis is an empirical method, it is important to assess the appropriateness of the data-collection plan. Good research practices for data collection associated with a conjoint analysis require the explanation and justification for the

sampling strategy, mode of administration, and an assessment of ethical considerations.

Sampling strategy

Sample-size calculations are particularly difficult for conjoint-analysis applications in health care. The appropriate sample size depends on the question format, the complexity of the choice tasks, the desired precision of the results, the degree of heterogeneity in the target population, the availability of respondents, and the need to conduct subgroup analyses [10]. Historically, researchers commonly applied rules of thumb, based on the number of attribute levels, to estimate sample size [58].

Orme [58] recommended sample sizes of at least 300 with a minimum of 200 respondents per group for subgroup analysis. Marshall et al. [59] reported that the mean sample size for conjoint-analysis studies in health care published between 2005 and 2008 was 259, with nearly 40% of the sample sizes in the range of 100 to 300 respondents.

Modes of administration

Conjoint-analysis surveys can be administered in many ways, including mail surveys using a paper-and-pencil survey instrument, non-mediated paper-and-pencil surveys completed at a finite set of study sites, electronic administration at a finite set of study sites using a laptop computer, or electronic administration over the Internet. The complexity of most conjoint-analysis questions probably precludes the use of telephone-based data collection, unless the survey instrument is mailed to respondents in advance. Interviewer-led administration of the survey may improve the quality of data because the interviewer can recognize that more explanation is needed, can more fully explain the task, and can answer questions (without leading the respondent).

Ethical considerations

Furthermore, those who design and conduct conjoint-analysis studies should consider the respondents and whether any issues would affect their ability to complete the survey. Some patient groups who have known difficulties with cognitive function—such as individuals with neurological diseases—may not be able to complete the tasks. In general, it is good practice to simplify the tasks as much as possible without compromising accuracy or completeness. Researchers should address the readability of the survey instrument and assess its appropriateness for the study population. Researchers also should address any ethical considerations or requirements mandated by ethics laws or practices or by ethics or institutional review boards.

Statistical analyses

Conjoint-analysis data and the modeling of preferences can require complex statistical analysis and modeling methods. There are several objectives when analyzing conjoint-analysis data. The primary objective is estimating the strength of preferences for the attributes and attribute levels included in the survey. Another objective might be estimating how preferences vary by individual respondent characteristics. For policy analysis, researchers may calculate how choice probabilities vary with changes in attributes or attribute levels or calculate secondary estimates of money equivalence (willingness to pay) [60], risk equivalence (maximum acceptable risk) [61], or time equivalence for various changes in attributes or attribute levels [31]. As part of the statistical analysis, good research practices require an assessment of the respondent characteristics, the quality of the responses, and a description and justification for the model estimation methods.

Respondent characteristics

The characteristics of the respondent sample should be reported and examined against the known characteristics of the population whose preferences researchers may want to generalize. Parametric and nonparametric statistical tests, such as the chi-square goodness-of-fit test, the Student's *t* test, or the Kolmogorov-Smirnov test, are available to test the hypothesis that the respondent sample has been drawn from the desired population. Furthermore, if data on the characteristics of those respondents who did not complete the survey in whole or in part are available, it is important to examine the differences between responders and non-responders or between other subsets of the respondent sample.

Quality of responses

The quality of responses can be assessed by evaluating the internal validity of the data. Examples include a repeated question, an alternative whose attribute levels are all better than the attributes of another alternative in the choice set, or three questions that support a check of preference transitivity. Analysis of these data can include a tabulation of response errors, how response errors relate to such demographic characteristics as age and education, and interacting dummy variables for response errors with attributes as one would do for other individual-specific characteristics.

Another type of check is to identify respondents who always or nearly always choose the alternative with the best level of one attribute. Preferences that are dominated by a single attribute can bias model estimation. For any failure of internal validity, it generally is better to include statistical controls for the problem rather than simply drop respondents from the data set.

Model estimation

Most conjoint analyses obtain multiple responses from each respondent. The data thus have the characteristics of cross-section panel data. Researchers should ensure that the statistical analysis of the data accounts for within-subject correlation. Thus, researchers who estimate these models should test that the data being analyzed are consistent with the assumptions for the model.

Researchers need to address how the independent variables are coded in the data. Attribute levels may be categorical (mild, moderate, severe), or continuous (1 day, 3 days, 7 days). Continuous variables can be modeled as either continuous or categorical. If continuous models are used, standard tests should be used to test whether the data are consistent with a linear, log, quadratic, or other functional form. Categorical models avoid imposing any functional form on preference weights and provide a validity check on the correct ordering of naturally ordered attribute levels. In addition, researchers need to consider whether categorical attribute levels should be specified as dummy variables or effects-coded variables. When effects coding is used, zero corresponds to the mean effect for each attribute, rather than the combination of all the omitted categories, and the parameter for the omitted category is the negative sum of the included-category parameters. Effects coding has desirable properties in modeling conjoint-analysis data and is widely used in many conjoint-analysis applications. However, effects coding is unfamiliar to most health care researchers and thus can complicate the presentation of the results.

Preference variation among individuals that is unaccounted for in modeling can result in biased estimates. Variations in preferences that arise from differences in individual characteristics such as age, education, gender, and health status often are of clinical or policy interest. This type of preference variation can be incorporated into the modeling process by interacting individual characteristics with the attributes included in the conjoint analysis. Researchers

also may consider split-sample analysis if sample sizes are sufficiently large.

A mixed-logit or random-parameter logit model allows for unobserved or random preference variation; it also can incorporate the cross-sectional panel structure of the data. The mixed-logit model increasingly is used for studies published in peer-reviewed journals. However, it can be difficult to implement because researchers are required to assume that the random-preference variation between respondents follows a particular pre-specified distribution. Further, mixed-logit estimation involves complicated statistical estimation techniques that can result in biased parameter estimates when the simulation methods fail to converge [23]. Latent-class models account for preference variation by using the data to identify groups of respondents with similar preferences [10] and may be preferable to mixed-logit models because researchers are not required to make assumptions about the distribution of preferences across respondents. Hierarchical Bayes is another approach to estimating preference variation. It directly estimates a different preference parameter for each respondent.

Results and conclusions

Researchers often are tempted to make inferences and predictions that go beyond what the data and methods can support. Evaluating the validity of results and conclusions requires consideration of the research question, as well as other aspects of the design and analysis. In order to assess the validity of the study, good research practices require an assessment of the study results and conclusions and a consideration of any study limitations.

Study results

The results should present the statistical findings in the context of the research question and should be presented in sufficient detail. The results should state which attributes or levels (and interaction terms, if relevant) were or were not significant, and the results should report uncertainty associated with estimates. Findings should be interpreted in the context of the choice being considered.

For example, in the multiple sclerosis example previously cited, the results could indicate that the rate of relapse was a significant attribute, and a negative coefficient could imply that higher rates of relapse were less preferred. If attributes and levels were found to be nonsignificant in the statistical analysis, these findings should be clearly stated in the results. Results also should provide interpretation of the relative value of specific attributes. For example, how the acceptable waiting time for nonemergency surgery varies with the rate of surgical errors (i.e., the marginal willingness-to-wait for a reduced rate[s] of surgical errors). Statistical uncertainty should be reported in a manner consistent with the type of model selected. If alternative model specifications were tested, the results of these alternative analyses should be described or presented in full.

Equivalence calculations, such as willingness to pay or maximum acceptable risk, require dividing utility differences that result from some change in efficacy or side effects by the incremental utility of one dollar (\$1.00) or by a 1% chance of a side effect. Confidence intervals always should be reported for such calculations and can be obtained by the delta or Krinsky-Robb method [62].

Study conclusions

The conclusions section should identify key findings of the study in the context of the original research question. A key element of any research study is to provide a relevant framework for interpreting the results: whether the results are consistent with or differ from existing studies in the literature and how this study ex-

tends existing research all should be clearly identified and discussed.

Study limitations

Limitations of the study and the potential effect(s) of these limitations on results should be clearly identified in the discussion section. Limitations can arise from selection of attributes and/or levels, such as simplifications adopted during survey development in order to produce a feasible design, possible correlation among selected attributes, or other design features (e.g., the inclusion or exclusion of an opt-out option). Assumptions underlying the analytic approach also may affect interpretation of results and should be discussed. If the study population is not representative of the population, this may limit generalizability of findings. Any extrapolation of results beyond the study population should be qualified and discussed.

Study presentation

Good research practices for the application of conjoint analysis in health care research not only require that the study is conducted well, but that it also is presented appropriately. This requires an adequate explanation of the research context, a description of the instruments and methods used, and a discussion of the study implications.

Research context

The study's importance and context must be adequately motivated so as to answer the "so what" question. Key background literature should be cited to place the study in an appropriate clinical or health policy context and to identify gaps in current knowledge that are important to researchers or decision makers. The specific contribution of the study, in terms of innovative methods or an important application, should be clearly stated at the end of the introduction.

The text describing the study should be worded and structured appropriately for the target journal and audience. Journals vary in both the type of reviewers and the eventual readers. In general, the use of jargon should be minimized. Acronyms and technical language (e.g., "importance weights" and "fractional factorial design") should be clearly defined, with any alternative terms included with the definition. A journal such as *Value in Health* has reviewers and readers who are familiar with conjoint-analysis methodology. This is unlikely in the case of a clinically focused journal.

Moreover, conjoint analysis is a relatively new area of research in health care. The use of technical terms is not always consistent among authors. For example, results may be referred to as "importance weights" or "preference weights." Such inconsistencies are confusing to reviewers and readers alike. Because there are no standardized rules for constructing a conjoint-analysis survey and because there are a large number of possible experimental designs, the methods and rationale for the study design must be adequately described. This includes the qualitative research conducted to identify the attributes and levels, the experimental design used to create the tasks, and the methods used to analyze the results. The matrix of attributes and levels and the final survey instrument should be submitted for review along with the paper.

Instrument and methods

A reviewer cannot provide a meaningful review of a conjoint-analysis paper without seeing the format and framing of the questions that generated the data. The properties of the experimental design should be described to provide a context for the strengths and limitations of the survey results. For example, if the experimental design does not allow interactions to be tested (a main-effects design), this assumption should be clearly disclosed in the methods section. Many journals will publish on their web-sites the da-

ta-collection instrument as a technical appendix to a conjoint-analysis manuscript. Even when the journal cannot or chooses not to publish the data-collection instrument, the data-collection instrument should be made available to reviewers and readers.

Study implications

Finally, the discussion section should focus on both the innovative features of the paper and the implications of the results for the target audience. The unique contributions of the study should be discussed and compared in the context of the current state of knowledge, as found in the published literature, and health care policy climate. However, as with all research, authors must be careful not to overstate the importance of their findings. Because conjoint analyses in health care are published in various types of journals and may use different terminology, it is important for authors to ensure that what may appear to be novel has not been conducted previously. In addition, it is important that authors inform readers that the results of a conjoint analysis often provide estimates of the value or importance of attributes to respondents, but that these results often do not predict future behavior or health outcomes.

As with all studies, the findings should be evaluated with respect to the research question that the study was designed to answer and the hypotheses to be tested in the study. If the target audience is a clinical one, the conclusions of the paper should focus on the clinical implications of the study findings. For example, the results can be translated into simple statements about their possible impact on physician practice. Alternatively, if a study was designed to inform health care policy, the findings about public, patient, or provider preferences can be translated into suggestions for increasing the appropriate use of health care services. For example, in a conjoint analysis of colorectal cancer screening tests, the findings were translated into changes in the rates of uptake of colorectal cancer screening based on the mix of alternative screening tests offered [63].

Conclusions

This report presents a checklist for good research practices of conjoint analysis in health care applications. The checklist was created from discussions and the experience of the task force members. This report presents researchers, reviewers, and readers with many questions to consider when assessing an application of conjoint analysis in health care. Though we have not aimed to identify best practices, we have intended to steer researchers away from bad practices. At this time, many unanswered questions related to the application of conjoint analysis in health care remain.

Conjoint analysis can be a powerful tool for quantifying decision makers' preferences for health care. There are, however, numerous approaches to conducting conjoint analysis and not all of them are appropriate for addressing every research question. In addition, approaches to conjoint analysis will continue to evolve as the number of empirical applications of conjoint analysis in health care increases and more researchers apply conjoint-analysis methods to a widening array of research questions. Therefore, researchers conducting conjoint analyses in health care should always be clear about the conjoint-analysis approaches they are using and why these approaches are appropriate to a particular study.

Acknowledgments

The Conjoint Analysis Checklist manuscript was initiated under the ISPOR Patient Reported Outcomes (PRO) & Patient Preferences Special Interest Group's Patient Preference Methods (PPM)—Conjoint Analysis Working Group. It was completed under the ISPOR

Conjoint Analysis in Health Good Research Practices Task Force. We are grateful for the participation and contributions of current and past members of the working group and task force. We are particularly indebted to Elizabeth Molsen, RN, and Marilyn Dix Smith, PhD, RPh, for challenging and empowering the working group and the subsequent Conjoint Analysis Good Research Practices Task Force to broaden the methods available to outcomes researchers worldwide. We are indebted to Christopher Carswell, Axel Mühlbacher, Bryan Orme, Liana Frankel, and Scott Grosse, all of whom served as external reviewers, as well as the many members of the ISPOR Patient Reported Outcomes (PRO) & Patient Preferences Special Interest Group who took the time to comment on the earlier draft version of this report.

REFERENCES

- [1] Chong C, Chen I, Naglie C, et al. Do clinical practice guidelines incorporate evidence on patient preferences? *Med Decis Making* 2007; 27:E63–4.
- [2] Khrahn M, Naglie G. The next step in guideline development. Incorporating patient preferences. *JAMA* 2008;300:436–8.
- [3] Marshall DA, Johnson FR, Kulin NA, et al. How do physician assessments of patient preferences for colorectal cancer screening tests differ from actual preferences? Comparison in Canada and the United States using a stated-choice survey. *Health Econ* 2009;18:1420–39.
- [4] Bridges J, Onukwugha E, Johnson FR, et al. Patient preference methods—a patient centered evaluation paradigm. *ISPOR Connections* 2007;13:4–7.
- [5] Bridges J. Stated-preference methods in health care evaluation: an emerging methodological paradigm in health economics. *Appl Health Econ Health Policy* 2003;2:213–24.
- [6] Bridges J, Kinter E, Kidane L, et al. Things are looking up since we started listening to patients: recent trends in the application of conjoint analysis in health 1970–2007. *Patient* 2008;1:273–82.
- [7] Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2:55–64.
- [8] Marshall DA, McGregor E, Currie G. Measuring preferences for colorectal cancer (CRC) screening—what are the implications for moving forward? *Patient* 2010;3:79–89.
- [9] Green PE, Srinivasan V. Conjoint analysis in consumer research: issues and outlook. *J Consum Res* 1978;5:103–23.
- [10] Louviere J, Hensher D, Swait J. *Stated Choice Methods: Analysis and Applications*. Cambridge, UK: Cambridge University Press, 2000.
- [11] Viney R, Lancsar E, Louviere J. Discrete choice experiments to measure consumer preferences for health and healthcare. *Expert Rev Pharmacoecon Outcomes Res* 2002;2:89–96.
- [12] Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics* 2008;26: 661–77.
- [13] Hensher DA, Rose JM, Greene WH. *Applied Choice Analysis: a Primer*. Cambridge, UK: Cambridge University Press, 2005.
- [14] Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000;320:1530–3.
- [15] Weston A, Fitzgerald P. Discrete choice experiment to derive willingness to pay for methyl aminolevulinic photodynamic therapy versus simple excision surgery in basal cell carcinoma. *Pharmacoeconomics* 2004;22:1195–208.
- [16] Mühlbacher AC, Lincke H-J, Nübling M. Evaluating patients' preferences for multiple myeloma therapy, a discrete-choice-experiment. *Psychosoc Med* 2008;5:Doc10.
- [17] Bridges J, Selck F, Gray G, et al. Common avoidance and determinants of demand for male circumcision in Johannesburg, South Africa. *Health Policy Plan* 2010; Oct 20 [Epub ahead of print].
- [18] Phillips KA, Maddala T, Johnson FR. Measuring preferences for health care interventions using conjoint analysis: an application to HIV testing. *Health Serv Res* 2002;37:1681–705.
- [19] Beusterien KM, Dziekan K, Flood E, et al. Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment. *Value Health* 2005;8:453–61.
- [20] Coast J, Salisbury C, de Berker D, et al. Preferences for aspects of a dermatology consultation. *Br J Dermatol* 2006;155:387–92.
- [21] King MT, Hall J, Lancsar E, et al. Patient preferences for managing asthma: results from a discrete choice experiment. *Health Econ* 2006; 16:703–17.
- [22] Peacock S, Apicella C, Andrews L, et al. A discrete choice experiment of preferences for genetic counseling among Jewish women seeking cancer genetics services. *Br J Cancer* 2006;95:1448–53.
- [23] Regier DA, Ryan M, Phimister E, et al. Bayesian and classical estimation of mixed logit: an application to genetic testing. *J Health Econ* 2009;28:598–610.
- [24] Roux L, Ubach C, Donaldson C, et al. Valuing the benefits of weight loss programs: an application of the discrete choice experiment. *Obesity Res Clin Pract* 2004;12:1342–51.
- [25] Hauber AB, Johnson FR, Sauriol L, et al. Risking health to avoid injections: preferences of Canadians with type 2 diabetes. *Diabetes Care* 2005;28:2243–5.
- [26] Johnson FR, Manjunath R, Mansfield CA, et al. High-risk individuals' willingness to pay for diabetes risk-reduction programs. *Diabetes Care* 2006;29:1351–56.
- [27] Wittink MN, Cary M, TenHave T, et al. Towards patient-centered care for depression: conjoint methods to tailor treatment based on preferences. *Patient* 2010;3:145–57.
- [28] Hauber AB, Johnson FR, Mohamed AF, et al. Older Americans' risk-benefit preferences for modifying the course of Alzheimer's disease. *Alzheimer Dis Assoc Disord* 2009;23:23–32.
- [29] Weighting and valuing quality adjusted life years: preliminary results from the Social Value of a QALY project. 2008. Available from: <http://www.ncl.ac.uk/ihs/research/publication/56274> [Accessed August 2, 2010].
- [30] Mohamed AF, Hauber AB, Johnson FR, et al. Patient preferences and linear scoring rules for patient reported outcomes. *Patient* 2010;3:217–27.
- [31] Johnson FR, Özdemir S, Mansfield CA, et al. Crohn's disease patients' benefit-risk preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology* 2007;133:769–79.
- [32] Bridges J, Searle S, Selck F, et al. Engaging families in the design of social marketing strategies for male circumcision services in Johannesburg, South Africa. *Soc Mar Q* 2010;16:60–76.
- [33] Opuni M, Bishai D, Gray GE, et al. Preferences for characteristics of antiretroviral therapy provision in Johannesburg, South Africa: results of a conjoint analysis. *AIDS Behav* 2010;14:807–15.
- [34] Fraenkel L. Conjoint analysis at the individual patient level: issues to consider as we move from a research to a clinical tool. *Patient* 2008;1: 251–3.
- [35] Nathan H, Bridges J, Schulick RD, et al. Understanding surgical decision-making in early hepatocellular carcinoma. *J Clin Oncol* 2011 Jan 4. [Epub ahead of print].
- [36] Shumway M. Preference weights for cost-outcome analyses of schizophrenia treatments: comparison of four stakeholder groups. *Schizophr Bull* 2003;29:257–66.
- [37] Peterson AM, Nau DP, Cramer JA, et al. A checklist for medication compliance and persistence studies using retrospective databases. *Value Health* 2007;10:3–12.
- [38] Wild D, Grove A, Martin M; ISPOR Task Force for Translation and Cultural Adaptation. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8:94–104.
- [39] Mauskopf JA, Sullivan SD, Annemans L, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on Good Research Practices—Budget Impact Analysis. *Value Health* 2007; 10:336–47.
- [40] Drummond MF, Sculpher MJ, Torrance GW, et al. *Methods for the Economics of Health Care Programmes (3rd ed.)*. New York: Oxford University Press, 2005.
- [41] Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. *J Health Serv Res Policy* 2007;12:25–30.
- [42] Kinter E, Schmeding A, Rudolph I, et al. Identifying patient-relevant endpoints among individuals with schizophrenia: an application of patient centered health technology assessment. *Int J Technol Assess Health Care* 2009;25:35–41.
- [43] Johnson FR, Mohamed AF, Ozdemir S, et al. How does cost matter in health-care discrete-choice experiments? *Health Econ* 2010 [Epub ahead of print].
- [44] Maddala T, Phillips KA, Johnson FR. An experiment on simplifying conjoint analysis designs for measuring preferences. *Health Econ* 2003;12:1035–47.
- [45] Huber J, Zwerina K. The importance of utility balance in efficient choice designs. *J Mar Res* 1996;33:307–17.
- [46] Kuhfeld WF. Marketing research methods in SAS. SAS Technical Paper MR2009. 2009. Available from: http://support.sas.com/resources/papers/tnote/tnote_marketresearch.html [Accessed August 2, 2010].
- [47] Kanninen B. Optimal design for multinomial choice experiments. *J Market Res* 2002;39:214–27.
- [48] Street DJ, Burgess L. *The Construction of Optimal Stated Choice Experiments*. New York: Wiley, 2007.

- [49] Chrzan K, Orme B. An overview and comparison of design strategies for choice-based conjoint analysis. Sawtooth Software: Research Paper Series. 2000. Available from: <http://www.sawtoothsoftware.com/download/techpap/desgncbc.pdf> [Accessed August 2, 2010].
- [50] Carlsson F, Martinsson P. Design techniques for stated-preference methods in health economics. *Health Econ* 2003;12:281–94.
- [51] Burgess L. Discrete choice experiments [computer software]. 2007. Department of Mathematical Sciences, University of Technology, Sydney, Australia. Available from: <http://crsu.science.uts.edu.au/choice/> [Accessed May 25, 2010].
- [52] Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling. *Expert Rev Pharmacoecon Outcomes Res* 2010;10:259–67.
- [53] Aday L. *Designing and Conducting Health Surveys: a Comprehensive Guide* (3rd ed.). San Francisco: Jossey-Bass, 2006.
- [54] Dillman DA, Smyth JD, Christian LM. *Internet, Mail, and Mixed Mode Surveys: the Tailored Design Method*. Hoboken, NJ: Wiley, 2008.
- [55] Lloyd AJ, McIntosh E, Williams AE, et al. How does patient's quality of life guide their preferences regarding aspects of asthma therapy? *Patient* 2008;1:309–16.
- [56] Lloyd AJ. Threats to the estimation of benefit: are preference elicitation methods accurate? *Health Econ* 2003;12:393–402.
- [57] Kjaer T, Bech M, Gyrd-Hansen D, et al. Ordering effect and price sensitivity in discrete choice experiments: need we worry? *Health Econ* 2006;15:1217–28.
- [58] Orme BK. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Madison, WI: Research Publishers LLC; 2006.
- [59] Marshall D, Bridges J, Hauber AB, et al. Conjoint analysis applications in health - how are studies being designed and reported? An update on current practice in the published literature between 2005 and 2008. *The Patient - Patient Centered Outcomes Research* 2010;3:249–56.
- [60] Kleinman L, McIntosh E, Ryan M, et al. Willingness to pay for complete symptom relief of gastroesophageal reflux disease. *Arch Intern Med* 2002;162:1361–6.
- [61] Johnson FR, Hauber AB, Özdemir S. Using conjoint analysis to estimate healthy-year equivalents for acute conditions: an application to vasomotor symptoms. *Value Health* 2009;12:146–52.
- [62] Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics* 2007;16: 827–40.
- [63] Marshall DA, Johnson FR, Phillips KA, et al. Measuring patient preferences for colorectal cancer screening using a choice-format survey. *Value Health* 2007;10:415–30.