available at www.sciencedirect.com

SCIENCE d DIRECT ®

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

ELSEVIER

EDITORIAL

Taking Conjoint Analysis to Task

The International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Conjoint Analysis Task Force was established to identify good research practices for applications of conjoint analysis (CA) in health. The report contains a discussion of issues raised when conducting such a study and presents a checklist for CA applications in health.

The ISPOR report begins by summarising applications of CA in health, noting the broad range of applications. For a recent review of such applications see de Bekker et al. [1]. Whilst CA was introduced into health to value characteristics beyond health outcomes, the technique now has a much broader use. de Bekker et al. [1] note the absence of CA within an economic evaluation framework. Quality-adjusted life years (QALYs) continue to be the preferred valuation method by bodies such as the National Institute for Health and Clinical Excellence. This is despite the increasing recognition of the importance of valuing the patient experience in the delivery of health care. Because CA can value patient experience aspects of care as well as health outcomes—and indeed the trade-offs between them—the increasing use of well conducted CA studies will hopefully lead to their increased use at the policy level. Thus, it is important that good practice is established when conducting CA studies. The checklist provided by the task force provides a useful guide to conducting a good CA study, as does the User Guide to conducting such studies published by Lancsar and Louviere [2].

As noted by the authors of the report, CA is an attribute-based measure of value. The key stages when conducting a study can be broadly defined as defining attributes and levels, determining choices to present to respondents (experimental design), development and administration of questionnaire, and analysis and interpretation of data. The task force discusses a number of issues across these stages when developing their checklist. Given the broad range of issues raised when applying CA, discussion of many of them is necessarily brief. The authors provide a useful list of references. Below I discuss some of the issues raised in a little more detail, highlighting additional points and useful reading. As pointed out by the task force, many of the issues raised in the report reflect the principles of a good survey design, and apply to all valuation methods.

Deriving attributes and levels is one of the most important stages when conducting a CA study. The most efficient experimental design and sophisticated form of econometric analysis cannot compensate if the attributes and levels are wrong. It is becoming increasingly important, when reporting attributes and levels, to report in detail how they were derived. Where qualitative work is employed, it is necessary to report a systematic and valid framework for analysis of such data. It is no longer sufficient to say, "Qualitative work was used to derive the attributes and levels, and they are shown in Table 1." The report notes that "If the number of possible attributes exceeds what one may find possible to pilot in a conjoint analysis . . .." It is worth noting here that this restriction on the number of attributes refers to an assumption of

the CA approach that individuals consider all the attributes, and make trade-offs. It is this assumption that allows the researcher to estimate the value of attributes; that is, how much money an individual would be willing to give up to have a lower waiting time. If "too many" attributes are included, there is concern that individuals will resort to simple decision-making strategies, such as always choose the cheapest option. Estimation of trade-offs, whilst possible at the analysis stage, would then not be valid.

When discussing possible approaches to deriving choice sets (experimental design methods), the authors note that resulting designs may include implausible combinations. They give the example of an orthogonal design, including a profile that combines severe nausea with no restrictions on activities of daily life. Methods exist for creating designs with restrictions, including nesting [2] and use of software packages such as SAS [3] and Ngene (http://www.choice-metrics.com/). This is an example of trading off respondent efficiency and statistical efficiency, as the authors note.

A few additional points that are not mentioned in the report are worth noting with regard to experimental design. In addition to the design methods mentioned, readily available designs exist [1]. A recent development in the design literature is to use prior assumptions about parameters (which can be obtained from pilot work) to improve efficiency. If the researcher wants to include interaction terms (i.e., preferences for one attribute depend on the levels of another attribute), this must be stated at the experimental design stage. The importance of testing for interaction terms is being increasingly recognised. For more on all these issues see de Bekker et al. [1].

The report considers the important issue of whether or not to include an opt-out or status-quo option within the choice sets (which would be added to the choices derived from the experimental design), stating, "The inclusion of an opt-out or status-quo option may be inappropriate for many types of research questions in health care." It is becoming increasingly recognized that for most CA health applications an opt-out/status-quo option is necessary. This is important when a cost attribute (or price proxy) is included, such that willingness to pay (a monetary measure of benefit) can be indirectly estimated (as is the case for many applications of CA in health). Failure to include an opt-out/status-quo option in such studies will misrepresent demand (because individuals may be forced to choose an option that is above their maximum willingness to pay). Forced choices may be more appropriate when individuals are being asked which options they prefer, as opposed to which options they would choose, and a price proxy is not included as an attribute. Examples may include using CA to estimate weights within the QALY framework, develop priority setting frameworks, and establish health professional preferences for treatment of patients [1]. When a researcher includes an opt-out or status-quo option, it is crucial to have information on the levels of the attributes for such options. Failure to have such information will result in the researcher not being able to analyse the opt-out /status-quo responses.

The authors further argue that '". . . the inclusion of an opt-out or status-quo option may have serious implications for the experimental design." I believe this statement is a bit strong. The authors are correct to note that the inclusion of the opt-out or status-quo option will compromise the statistical design properties. Where levels for the opt out/status quo are known in advance of data collection the implications for the efficiency of the design can be estimated in advance. It is also advisable to simulate response data and test that the desired model can be estimated. Where the opt-out levels are constant across respondents, the affect on the design efficiency is smaller. Where information is collected from individuals within the study on their current situation/status quo, and this varies across respondents, the effect on the statistical efficiency of the design is likely to be greater and cannot be estimated in advance of data collection. However, this is another example of where respondent efficiency will have been gained (because the choices presented are more realistic) at the expense of statistical efficiency, and the benefits are likely to outweigh the costs.

The report discusses response rates and incentives. The authors argue that "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." To date very few CA studies have offered incentives for completion. Within a publicly provided health care system the incentive is often to inform respondents that information on their preferences will influence the provision of health care. Personal experience, however, suggests that response rates to CA studies are falling, and methods to increase response rates do need to be considered. Incentives may therefore be considered, and research should test the effectiveness of different incentives.

The authors note that researchers need to address how the independent variables are coded, distinguishing between categorical and continuous variables. It is worth noting that estimation of trade-offs can only be conducted when a continuous variable is modelled. This continuous variable is most commonly price, but risk and time have also been used in health applications of CA [1]. Furthermore, for discussion of the pros and cons of effects coding versus dummy variable coding, the reader should see Bech and Gyrd Hansen [4] and for an application of effects coding see Watson et al. [5].

Conditional logit continues to be the most common method for analyzing CA response data [2]. The report identifies a number of more sophisticated analytical approaches, including mixed-logit and latent class, which drop some of the assumptions of the conditional logit model and allow for unobserved preference heterogeneity. I often wonder how useful such information is to policy makers because it is not possible to identify where preferences differ. A more useful approach may be to use conditional logit to gain better insight into observed variation that can potentially be acted upon by policy makers. The authors also mention Hierarchical Bayes for estimating preference parameters for each respondent. This approach is applicable when individuals are presented with all choices derived from the experimental design, although it does raise the issue of the extent to which policy can be developed at the individual level. However, when a blocked design is used, and individuals are presented with a fraction of the fractional factorial design, this approach is not appropriate.

Because CA relies on responses to hypothetical choices, it is important to check the validity of responses. Given the lack of a market for health care, tests of internal validity dominate. The report notes that such tests include repeated choices, dominant choices, transitivity tests, and nontrading responses (always choosing the alternative with the best level of a given attribute). Sen's expansion and contraction properties are also beginning to be applied and it is common practice to check if the signs of estimated parameters are consistent with a priori expectations [1,2]. The authors note that it is generally better to include statistical controls to allow for failures (rather than simply dropping respondents). It is also worth noting that qualitative work has found that individuals defined as failing such tests had "rational" reasons for doing so [1]. Lancsar and Louviere [2] note that random utility models are robust to violations of compensatory decision making and errors made by individuals in forming and revealing preferences, supporting the authors of this report.

I argue that the greatest challenge facing practitioners of CA is to test the external validity of responses; that is, the extent to which respondents behave in reality as they state in a hypothetical context. There has been very little research in this area [1], probably reflecting the difficulty of investigating this question in a publicly provided health care system where individuals have limited choice and do not pay at the point of consumption. Possible areas for future research include field experiments comparing revealed and stated choices for private health care goods, as well laboratory and classroom experiments [1]. Such research would help provide further guidance on good practice, and lend more credibility and confidence to results from well-conducted CA studies that follow good research practice (as described in this report and Lancsar and Louviere [2]). This would hopefully lead to the increased use of CA at the policy level, challenging the QALY as the main valuation method within economic evaluations.

Mandy Ryan, PhD
Health Economics Research Unit, University of Aberdeen, UK.

REFERENCES

[1] de Bekker-Grob E, Ryan M, Gerard M. Discrete choice experiments in health economics: a review of the literature. Health Econ. In press.

[2] Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 208;26:661–77.

[3] Kuhfeld W. Marketing Research Methods in the SAS System, Version 8 Edition. Cary, NY: SAS Institute, 2000.

[4] Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. Health Econ 2005;14:1079–83.

[5] Watson V, Ryan M, Watson E. Valuing experience factors in the provision of Chlamydia screening: an application to women attending the family planning clinic. Value Health 2009;12:621–3.