

Medical Device Decision Analysis: A Risk-Tolerance Pilot Study
0910-NEW
SUPPORTING STATEMENT Part B

B. Statistical Methods

Descriptions of Statistical Methods

Choice-format conjoint-analysis studies, sometimes called discrete-choice experiments, are designed specifically to provide information about individuals' willingness to accept tradeoffs among features of multi-attribute products. Choice-format conjoint analysis is based on the hedonic principle that products have various features, the attractiveness of a product to users depends on users' relative preferences for these features, and users are willing to accept tradeoffs among them. For example, some patients may be willing to tolerate mild-to-moderate treatment side effects to achieve greater weight loss, while other patients may not. Numerous studies have demonstrated variations in preferences and willingness to accept tradeoffs among attributes of medical interventions.¹

The choice-format conjoint survey instrument requires constructing a series of tradeoff questions that evaluate hypothetical weight-loss devices. Each hypothetical device consists of combinations of attribute levels for amount of weight loss obtained with the device, duration of weight loss, type of surgery required to implant the weight-loss device, diet restrictions associated with the weight-loss device, duration of side effects, chance of serious side effects, improvements in weight-related diseases, and risk of death associated with the weight-loss device. The combination of attributes and levels that respondents evaluate in a conjoint survey is known as the experimental design. These combinations must have statistical properties that allow estimating the preference weights of interest. The Contractor uses the SAS implementation of a commonly used D-optimal algorithm to construct a fractional factorial experimental design (Kuhfeld, 2010; Kuhfeld et al., 1994). The experimental design includes an appropriate number of pairs so that both main effects and possible selected interaction effects can be estimated. Main effects include the preference weight for each attribute level independent of the other attributes and levels included in the study. Interaction effects allow us to test whether the effect of one attribute level is independent of the other attributes and levels in the study.

Using choices between the constructed weight-loss devices, it is possible to estimate relative preference weights for each attribute level. Conjoint-analysis questions generate cross-section/time-series data that require analysis using advanced statistical techniques. The Contractor will use random-parameters logit (RPL)² to analyze the choice-format conjoint data collected in this survey. Unobserved variation in preferences across the sample can bias estimates in conventional choice models. RPL avoids this potential bias by estimating a distribution of preferences around each model parameter that accounts for variations among individual preferences not accounted for by the variables in the model.

¹ Attributes are generic product characteristics such as chance that the treatment will work well. Chance that the treatment will work well may take on several possible levels such as works well in 25% of patients, works well in 75% of patients, and works well in 100% of patients.

² RPL is also referred to as mixed logit, random coefficients logit, and error-components logit.

The flexible correlation structure of RPL also accounts for within-sample correlation in the randomized question sequence for each respondent.

The results of this analysis will include the following:

- Model log-odds (relative preference weight) estimates
- Odds ratio tests of whether selected treatment-profile preferences are significantly different from a specified standard-of-care or current-treatment profile.

Once these results have been estimated, they can be used to calculate the following:

- Predicted choice shares for two or more treatments with specified attribute levels, indicating the predicted proportion of respondents who would choose each treatment profile
- Maximum acceptable side-effect risk for selected improvements in efficacy. Other possible measures of risk tolerance include minimum acceptable efficacy for given side-effect risks and incremental net benefits.

1. Respondent Universe and Sampling Methods

The respondent universe includes adults with a self-reported body-mass index (BMI) during the last 3 years of 30 kg/m² or above. Among these subjects, are those who have had a weight-reduction procedure such as gastric bypass and gastric banding. Respondents must also be able to provide informed consent and read and understand English.

It is expected that the Contractor will perform several activities intended to ensure the final data set contains a statistically valid number of respondents to be able to infer the survey result to the sample universe, e.g., 450 respondents. These procedures are described below.

A 1,000-person sample with a self-reported BMI during the last 3 years of 30 kg/m² or above (includes no more than one panelist per household) will be selected from approximately 11,642 panel members, from the KnowledgePanel®, a national online panel owned by GfK Knowledge Networks (KN). KnowledgePanel® is the only online panel that is representative of the U.S. population, providing sampling coverage of 97% of the U.S. adult population through address-based sampling. Because every sample unit has a known selection probability, KnowledgePanel® is not susceptible to the “professional respondent” problem and other hazards of “opt-in” online panels based on convenience sampling. Recent comparison research has demonstrated that KnowledgePanel®’s accuracy rates are comparable to high-quality random-digit dialing surveys and superior to online opt-in panels (Yeager et al., 2009).

Our data-collection contractor will produce and deliver statistical sample weights incorporating the probabilities of selection and modified by post-stratification weighting based on population benchmarks from the Current Population Survey or a similar

benchmark source. The panel sample weights are adjusted to demographic benchmarks to reduce bias due to nonresponse and other nonsampling errors.

Sampled panel members will receive an email inviting them to participate in the study (Attachment B). We estimate that approximately 700 panel members will choose to participate, and complete the introduction and consent form (included in Attachment A). We estimate 450 respondents will complete the full survey. Among these subjects, approximately 150 also will have had a weight-reduction procedure such as gastric bypass and gastric banding.

Appropriate sample size depends on a number of criteria, including the question format, the complexity of the choice task, the desired precision of the results, and the need to conduct subgroup analyses (Louviere et al., 2000). Researchers commonly apply a rule of thumb such as that proposed by Orme (2006). A choice-format conjoint study design with 6 to 8 attributes, each with three levels, and 8 to 10 choice questions per respondent requires at least 200 to 300 respondents to estimate a preference model with acceptable confidence intervals for all parameters. Sample size larger than 300 improves the representativeness of the sample. The sample size proposed should provide sufficient power to detect significant differences between the three device label formats to be tested and the types of participants to be recruited. All estimates will be reported with 95% confidence intervals.

2. Procedures for the Collection of Information

The survey sample will be drawn from eligible members of the KN panel by using an implicitly stratified systematic sample design based on the methodology for which KN was assigned a U.S. Patent (U.S. Patent No. 7,269,570) in September 2007. The selection methodology, which has been used by KN since 2000, ensures that KN panel samples will closely track the U.S. population and survey panelists will not be overburdened with survey requests.

Once sampled for a survey, panel members receive a personal notification email on their computer (Attachment B) letting them know there is a new survey available for them to take. The email notification contains a button to start the survey. Alternatively, panel members can access the online survey by logging into their specific panel home page, where they will find a hyperlink to surveys for which they have been selected.

All survey results will be collated and analyzed by RTI-HS. The results of these data collection activities will provide vital information to the FDA experts.

3. Methods to Maximize Response Rates and Deal with Non-Response

Nonresponse bias will be analyzed by comparing the ancillary data available for the sample invited to participate in the study but did not complete the survey against the subset of recruited study participants that completed the survey. Statistically significant

differences in marginal distributions of person-level and household-level characteristics would indicate nonresponse bias relative to the invited sample. Statistical comparisons for specific studies can be made between the total invited sample from the panel and the estimating sample for the characteristics noted above (e.g., categories of age, education, race, ethnicity, gender, head of household status, household size, housing type, income, marital status, metropolitan residence, home ownership, state, employment, and internet access). An aggregate error rate can be calculated as the sum of the differences in the distributions between the expected values from the total invited sample compared to the actual values (from the estimating sample of completed surveys).

We expect that the maximum time for reporting burden for any given survey would not exceed 30 minutes, but most respondents will complete in fewer than 25 minutes. Also, as described below, we have reduced the number of questions each respondent will answer.

The conjoint survey instrument requires assembling a series of choice questions and evaluating hypothetical treatments. Each hypothetical treatment consists of combinations of treatment attributes. These combinations must have statistical properties that allow estimation of the preference weights of interest. The validity of resulting estimates depends on subjects' successes in completing the trade-off tasks.

Most choice-format conjoint applications currently use a D-optimal design to reduce the number of paired comparisons to the smallest number necessary for efficient estimation of preference weights (Dey, 1985; Huber and Zwerina, 1996; Kanninen, 2002; Kuhfeld et al., 1994). Efficient designs can be produced using an iterative computer algorithm (Zwerina et al., 1996). We use a variation of a commonly used D-optimal algorithm to construct a fractional factorial experimental design. The resulting experimental design included 120 treatment pairs.

Because there is a limit to the number of choice questions each respondent can reasonably answer before becoming fatigued, we will split the experimental design into 15 blocks. Each block will have 8 questions, and questions will not be repeated across blocks. Respondents will see only one block when answering the survey. This will greatly reduce the number of questions that respondents answer and therefore should reduce measurement error in the data while ensuring that the all questions in the full design are presented to subjects in the sample. Each subject will be assigned a block randomly.

Studies (Schwappach and Strasmann, 2005; Maddala et al., 2003; Johnson and Desvousges, 1997) suggest that subjects' learning during the first choice questions and subjects' fatigue after answering many choice questions contribute to measurement error and thus affect preferences estimates. To avoid having some questions regularly affected by learning and fatigue, the order of the choice questions in each block will be randomized for each respondent.

4. Test of Procedures or Methods to be Undertaken

RTI-HS conducted 9 face-to-face pretest interviews between February 13, 2012 and March 6, 2012, in Raleigh, North Carolina. Respondents were invited to participate in the pretests if they had a calculated body mass index of 30 or above, with at least three respondents having prior experience with bariatric surgery or gastric banding.

During the pretest interviews, respondents were asked to think aloud as they completed the survey. After completing the survey in this manner, respondents were asked a series of debriefing questions to determine whether they understood the definitions and instructions, accepted the hypothetical context of the survey, and successfully completed the trade-off questions in the conjoint survey instrument as instructed. Before and during the pretest interviews, RTI-HS identified specific issues that could negatively affect the quality of the preference information collected with the survey instrument and made changes to the instrument to address these issues.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The information will be collected and analyzed by RTI Health Solutions, the contractor for the information collection.

REFERENCES

- Kuhfeld W, Tobias F, Garratt M. Efficient experimental design with marketing research applications. *J Market Res* 1994;31:545-57.
- Kuhfeld W. *Marketing research methods in SAS: experimental design, choice, conjoint, and graphical techniques*. Cary, NC: SAS Institute Inc.; 2010.
- Dey A. *Orthogonal fractional factorial designs*. New York, NY: Halstead Press; 1985.
- Huber J, Zwerina K. The importance of utility balance in efficient choice designs. *J Marketing Res* 1996;33:307-17.
- Kanninen B. Optimal design for multinomial choice experiments. *J Market Res* 2002;39:214-27.
- Zwerina K, Huber J, Kuhfeld W. *A general method for constructing efficient choice designs*. Durham, NC: Fuqua School of Business, Duke University; 1996.
- Schwappach DLB, Strasmann TJ. Quick and dirty numbers? The reliability of a stated-preference technique for the measurement of preferences for resource allocation. *J Health Econ* 2005;25(3):432-48.
- Maddala T, Philips KA, Johnson FR. An experiment simplifying conjoint analysis designs for measuring preferences. *Health Econ* 2003;12(12):1035-47.
- Johnson FR, Desvousges WH. Estimating stated preferences with rated-pair data: environmental, health, and employment effects of energy programs. *J Environ Econ Manag* 1997;34:79-99.
- DiSogra CJ, Dennis JM, Fahimi M. On the quality of ancillary data available for address-based sampling. *Conference Proceedings of the 2010 Joint Statistical Meetings and in review*.
- Louviere JJ, Hensher DA, Swait JD. *Stated choice methods: analysis and applications*. New York: Cambridge University Press; 2000.
- Orme BK. *Getting started with conjoint analysis: strategies for product design and pricing research*. Madison, WI: Research Publishers LLC; 2006.
- Yeager D, Krosnick J, Chang LC, Harold J, Levindusky M, Simpser A, and Wang R (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly*, 75 (4), 709-747.