

## BACKGROUND

- Quality-of-life instruments are routinely validated. What constitutes validation of a conjoint or stated-choice (SC) survey instrument?
- Utility theory requires that stated choices conform to the axioms of choice: completeness, reflexivity, transitivity, continuity and monotonicity of preferences.
- Welfare measures from SC data are internally valid if preferences conform to the axioms of choice listed above.
- Internal-validity tests can be incorporated into any SC survey to check whether preferences conform to one choice axiom or another.
- Failure of an internal-validity test is defined as a violation of the corresponding choice axiom.
- Data on internal-validity failures from a specific survey are typically used for testing whether the preference estimates regarding the object of that survey are affected.
- Data on internal-validity failures can also be pooled across surveys to explore the general determinants of such failures, since the choice axioms are universal.
- Several SC studies investigated internal validity. They either dropped the subjects who failed internal-validity tests or reported two sets of estimates.<sup>1-5</sup>
- None of the SC studies that included internal-validity tests reported the characteristics of respondents who were more likely to fail them.
- We are aware of no study that investigated the internal validity of SC preferences by pooling responses across several different SC studies.

## OBJECTIVE

To estimate the extent to which internal-validity test outcomes in health care SC surveys are explained by personal characteristics of the respondents, controlling for survey complexity.

## METHODS

### Relevant Terms

An SC survey presents the respondent with a sequence of trade-off tasks, or choice sets, with two alternatives. The alternatives have different levels or values of a common list of features, or attributes. On each task the respondent chooses alternative A or alternative B by comparing their relative levels for all the attributes considered jointly.

### Internal-Validity Tests

- Within-set monotonicity:** Alternative A is at least as good as alternative B for all attributes.

Alternative A is better than alternative B for at least one attribute.

**Alternative A must be preferred to alternative B.**

- Cross-set monotonicity:** Alternative A is preferred to alternative B.

Alternative B is at least as good as alternative C for all attributes.

**Alternative A must be preferred to alternative C.**

- Stability:** Alternative A is preferred to alternative B at one point in the trade-off sequence.

**Alternative A must be preferred to alternative B when same question is repeated.**

- Transitivity:** Alternative A is preferred to alternative B.

Alternative B is preferred to alternative C.

**Alternative A must be preferred to alternative C.**

## METHODS (continued)

### Subjects and Therapeutic Areas

- US resident adult patients (18 years or older)
- Eight different therapeutic areas
  - Crohn's disease
  - Bipolar disorder
  - Osteoarthritis
  - Rheumatoid arthritis
  - Multiple sclerosis
  - Type II diabetes
  - Colorectal cancer screening
  - HIV screening

### Data and Analysis

- All data collected using Web-enabled survey instruments.
- Subjects completed 8 to 12 conjoint trade-off tasks (example in Figure 1).
- Every survey included at least one internal-validity test.
- General linear model (GLM) with a binomial error structure:
  - Dependent variable: ratio of test failures to total number of tests per subject
  - Explanatory variables: personal characteristics (gender, race, education, age, annual household income)
- Probit models
  - Indicator variable: pass/fail outcome for each internal-validity test
  - Explanatory variables: personal characteristics (gender, race, education, age, annual household income) and total number of attribute levels as a measure of survey complexity<sup>6-8</sup>

Figure 1. Example of Conjoint Trade-off Task for Osteoarthritis Treatment Option

Feature	Treatment A	Treatment B
Pain	No Pain Extreme Pain	No Pain Extreme Pain
Stiffness	No Stiffness Extreme Stiffness	No Stiffness Extreme Stiffness
Stomach Problems	Occasional mild symptoms. Treat with over-the-counter medicines.	Frequent moderate symptoms. Treat with a prescription medicine.
Risk of Bleeding Ulcer	10 patients out of 1,000 (1%) will have a bleeding ulcer	None will have a bleeding ulcer
Risk of Heart Attack	None will have a heart attack	10 patients out of 1,000 (1%) will have a heart attack

Which treatment would you choose (check one)?

Treatment A  Treatment B

## ESTIMATION RESULTS (Table 1)

### All-Test Model (GLM model in column 1)

- Higher income, more education, being white and being female significantly decrease the failure rate.
- Survey complexity significantly increases the failure rate.
- Older respondents perform as well as younger respondents.

### Separate-Test Models (probit models in columns 2 through 5)

- The same general findings with mixed statistical significance, except for transitivity.

Table 1. Estimation Results

	(1)	(2)	(3)	(4)	(5)
	Failure rate (all tests)	Within-set monotonicity	Cross-set monotonicity	Stability	Transitivity
Male	0.091	0.119	0.047	0.085	-0.262
White	-0.156	-0.282	-0.352	-0.112	0.378
Natural log of years of education	-0.441	-0.990	-0.064	-0.335	-0.349
Natural log of years of age	0.059	0.145	-0.042	-0.055	0.210
Natural log of income	-0.101	-0.085	-0.259	-0.058	-0.047
Number of levels	0.035	0.112	0.021	0.024	0.040
Constant	-0.115	-0.780	1.126	0.314	-2.342
Observations	3,639	2,139	478	2,547	1,342

Highlighted: significant at 1% or 5%

## REFERENCES

- Carlsson F, Matinsson P. J Environ Econ Manage 2001;41:179-92.
- Ryan M, Bate A. Appl Econ Lett 2001;8:59-63.
- Miguel FS, Ryan M, Scott A. J Econ Behav Org 2002;48:1-14.
- Bryan S, Gold L, Sheldon R, Buxton M. Health Econ 2000;9:385-95.
- Schwappach DLB, Strasmann, TJ. J Health Econ 2005;25(3):432-48.
- Johnson FR, Banzhaf M, Desvousges W. (2000). Health Econ 9:295-317.
- Saelensminde, K. Environ Resour Econ 2002;23:403-20.
- DeShazo JR, Fermo G. J Environ Econ Manage 2002;44:123-43.

## CONCLUSIONS

- Utility-theoretic tests
  - Provide evidence on the validity of preference estimates.
  - Identify surveys that may be too complicated, too long, or otherwise flawed.
  - Indicate how well the instrument works for particular groups of subjects.
- Detecting inattentive respondents with internal-validity tests has two disadvantages:
  - Respondents are detected post-recruitment.
  - Most internal-validity tests reduce the statistical efficiency of the experimental design.
- High failure rates may suggest need for
  - More careful pretesting to identify reasons for poor performance.
  - Reducing the number of attributes, increasing number of attribute overlaps in each choice question, or reducing number of choice questions asked.
  - In-person interviews or higher incentive payments to increase the attention respondents give to the survey.

## CONTACT INFORMATION

**Gavril Huiber, PhD**  
Research Director  
RTI Health Solutions, RTI International  
3040 Cornwallis Road, PO Box 12194  
Research Triangle Park, NC 27709-2194  
Phone: 919.541.6590  
Fax: 919.541.7222  
E-mail: huiber@rti.org  
Presented at:  
12th Annual International Meeting of the  
International Society for Pharmacoeconomics  
and Outcomes Research  
May 19-23, 2007  
Arlington, VA