Quantifying Women's Stated Benefit–Risk Trade-Off Preferences for IBS Treatment Outcomes

F. Reed Johnson, PhD,[1] A. Brett Hauber, PhD,[1] Semra Özdemir, MS,[2] Larry Lynd, PhD[3]

[1]RTI International, Research Triangle Park, NC, USA; [2]University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [3]University of British Columbia, Vancouver, BC, Canada

ABSTRACT

Background: The Food and Drug Administration, currently, is exploring quantitative benefit–risk methods to support regulatory decision-making. A scientifically valid method for assessing patients' benefit–risk trade-off preferences is needed to compare risks and benefits in a common metric. **Objectives:** The study aims to quantify the maximum acceptable risk (MAR) of treatment-related adverse events (AEs) that women with diarrhea-predominant irritable bowel syndrome (IBS) are willing to accept in exchange for symptom relief. **Methods:** Research design: A stated-choice survey was used to elicit trade-off preferences among constructed treatment profiles, each defined by symptom severity and treatment-related AEs. Symptom attributes included frequency of abdominal pain and discomfort, frequency of diarrhea, and frequency of urgency. AE attributes included frequency of mild-to-moderate constipation and the risk of four possible serious AEs. Subjects: A Web-enabled survey was administered to 589 female US residents at least 18 years of age with a self-reported diagnosis of diarrhea-predominant IBS. **Measures:** Preference weights and MAR were estimated using mixed-logit methods. **Results:** Subjects were willing to accept higher risks of serious AEs in return for treatments offering better symptom control. For an improvement from the lowest to the highest of four benefit levels, subjects were willing to tolerate a 2.65% increase in impacted-bowel risk, but only a 1.34% increase in perforated-bowel risk. **Conclusions:** Variation in MARs across AE types is consistent with the relative seriousness of the AEs. Stated-preference methods offer a scientifically valid approach to quantifying benefit–risk trade-off preferences that can be used to inform regulatory decision-making. *Keywords:* benefit–risk analysis, conjoint analysis, incremental net benefits, irritable bowel syndrome, maximum acceptable risk.

Introduction

Several recent and well-publicized events involving withdrawals of drugs from the US market have highlighted the problem of balancing benefits and risks [1]. In all these cases, interventions offering potentially significant therapeutic benefits were found to carry increased risks of serious and, possibly, life-threatening adverse events (AEs). Decisions to halt the development or marketing of such therapies clearly require balancing benefits and risks. Despite the importance of establishing consistent and principled criteria for determining when benefits outweigh risks, experts have provided surprisingly little guidance to help decision-makers evaluate such trade-offs.

A review of past examples of product withdrawals and risk-management decisions in different countries reported that decisions, often, are inconsistent and are based on very limited scientific evidence beyond the original clinical trial data relating to safety and efficacy [2]. In addition, the recent Institute of Medicine report, The Future of Drug Safety, a study requested by the US Food and Drug Administration (FDA) to address recognized shortcomings of the US drug-safety system, noted that "in both the preapproval and the postmarketing setting, the risk-benefit analysis that currently goes into regulatory decisions appears to be ad hoc, informal, and qualitative" [3]. The FDA Amendments Act of 2007 called on the agency to collaborate with public and private entities to improve the quality of benefit–risk analysis (H.R. 3580 [Public Law 110-85] §904).

Regulatory agencies do not require quantifying or even formal consideration of the values of patients, physicians, or other stakeholders in risk evaluations. The values and risk tolerance of patients with a particular condition may be presented to advisory panels and policymakers either individually or through advocacy organizations; however, there is no transparent or consistent mechanism currently in place for quantifying systematically the values and risk tolerance of these ultimate stakeholders.

The case of alosetron illustrates the need for quantitative, preference-based, benefit–risk analysis. Alosetron was approved for marketing by the FDA in February 2000. The approved indication was for diarrhea-predominant irritable bowel syndrome (IBS) in women only. Although clinical trials demonstrated that alosetron provided relief of abdominal pain and discomfort, improvement in urgency, and decreased frequency of diarrhea [4], safety signals indicated the possibility of serious gastrointestinal AEs. The most serious risk of concern associated with alosetron was the possibility that women with IBS taking the drug would develop a perforated bowel requiring surgery. As a result, alosetron was withdrawn from the market 9 months after launch. In June 2002, in response to pressure from patient organizations and reanalysis of data, the FDA reapproved the drug for restricted use in a more targeted indication under a risk-management program.

Understanding the value that women with IBS place on treatment outcomes and their willingness to accept risks in return for treatment benefits can help inform future regulatory and risk-management decision-making. In this study, we employed well-established stated-choice (SC) methods (also known as choice-format conjoint analysis or discrete-choice experiments) to quantify the maximum acceptable risk (MAR) of treatment-related AEs that women with diarrhea-predominant IBS are willing to accept in exchange for symptom relief. In a related

study published in this journal, these estimates were used to construct preference weights, which were used in an event-simulation model, to estimate the incremental net benefits of alosetron [5].

Methods

SC Methods

In an SC survey, a sample of patients, physicians, or caregivers are asked to choose between treatment options where attribute levels are varied across options and across choice tasks [6–8]. SC methods yield quantitative estimates of trade-offs subjects are willing to make among treatment attributes and yield estimates of relative preference weights. These weights can be used to populate models in lieu of conventional health-state utilities or to scale therapeutic improvements in terms of one of the attributes, such as money, risk, or time [9,10].

Survey Development

Treatment-related benefits and risks were identified from a review of the literature, consultations with medical experts involved in IBS clinical trials, and interviews with women with diarrhea-predominant IBS. In each treatment-choice question in the survey, symptom attributes included 1) frequency of abdominal pain and discomfort; 2) frequency of diarrhea; and 3) frequency of feelings of urgency. AE attributes included 1) frequency of mild-to-moderate constipation and 2) risks of four additional AEs—three of them serious. AE risks included probabilities of moderate colitis, impacted bowel, severe colitis, and perforated bowel (Table 1). Probabilities of experiencing each AE ranged from 0% to 1%. Pretests were conducted using in-person interviews with eight women between 29 years and 60 years of age with a self-reported diagnosis of diarrhea-predominant IBS.

Figure 1 provides an example of the SC question format. We employed a commonly used algorithm to construct a statistically efficient experimental design resulting in 48 treatment-choice pairs [11–15]. We implemented an extension of Zwerina et al.'s algorithm that searches for maximum D-efficiency, subject to no dominated pairs, minimal overlaps, and best level balance [13,15]. The design with the highest D-score achieved an acceptable level of statistical efficiency for our sample size, as indicated by confidence intervals on the parameter estimates. Kanninen shows that prior information on parameter values can be used to improve design efficiency [14]. We did not have any information with which to specify priors for the parameters other than natural ordering, which we used in the search algorithm to screen out dominated pairs.

To reduce cognitive and time burden, treatment-choice questions were blocked into six sets of eight questions, and each subject was randomly assigned to one of the six sets. The final survey instrument also included questions regarding each subject's personal characteristics (e.g., age and education) and experience with IBS and IBS treatments. The survey was approved by the Research Triangle Institute's Office of Research Protection and Ethics.

Survey Sample

The Web-enabled survey was programed by Ipsos Observer, an international survey-research firm [16], and administered to female members of the Ipsos Online Access Panel. All subjects were required to have had a physician diagnosis of diarrhea-predominant IBS (self-reported) and to be US residents, at least

Table 1 Irritable bowel syndrome (IBS) treatment attributes and levels

Treatment attribute	Levels
Frequency of abdominal pain and discomfort	No IBS pain and discomfort
	IBS pain and discomfort for 1 week a month
	IBS pain and discomfort for 2 weeks a month
	IBS pain and discomfort for 3 weeks a month
	IBS pain and discomfort for 4 weeks a month
Diarrhea frequency	No diarrhea
	Diarrhea 2 times a day
	Diarrhea 4 times a day
	Diarrhea more than 4 times a day
Urgency frequency	No urgency
	Urgency 2 days a week
	Urgency 5 days a week
	Urgency 7 days a week
Frequency of mild-to-moderate constipation	No constipation
	Constipation 1 week a month
	Constipation 2 weeks a month
	Constipation 3 weeks a month
	Constipation 4 weeks a month
Chance of serious adverse event	No chance of severe adverse event
	1 person out of 1000 (0.1%) will have moderate colitis requiring doctor's care
	5 people out of 1000 (0.5%) will have moderate colitis requiring doctor's care
	10 people out of 1000 (1%) will have moderate colitis requiring doctor's care
	1 person out of 1000 (0.1%) will have an impacted bowel requiring doctor's care
	5 people out of 1000 (0.5%) will have an impacted bowel requiring doctor's care
	10 people out of 1000 (1%) will have an impacted bowel requiring doctor's care
	1 person out of 1000 (0.1%) will have severe colitis requiring hospitalization
	5 people out of 1000 (0.5%) will have severe colitis requiring hospitalization
	10 people out of 1000 (1%) will have severe colitis requiring hospitalization
	1 person out of 1000 (0.1%) will have a perforated bowel requiring surgery
	5 people out of 1000 (0.5%) will have a perforated bowel requiring surgery
	10 people out of 1000 (1%) will have a perforated bowel requiring surgery

18 years of age. Study subjects were entered into a drawing to win one of the five $100 cash prizes offered as an incentive for their participation.

Statistical Analysis

We used multivariate, random-parameters panel-logit regression to estimate preference parameters for each attribute level [17]. Explanatory variables in the random-parameters logit model included all attribute levels listed in Table 1. All statistical analyses were conducted using GAUSS 7.0 (Aptech Systems, Inc., Black Diamond, WA) [18].

The parameter estimates from SC models are preference weights that indicate the relative strength of subjects' preference for each attribute level. Attribute levels were effects coded. The preference weight for the omitted category is the negative sum of the included-category parameters [19]. Thus, zero is the mean effect for each attribute, and positive and negative preference weights are interpreted relative to the mean effect of the attribute on treatment choice.

MAR Calculations

Estimated preference parameters were used to calculate the mean MAR for each serious AE—impacted bowel, severe colitis, and
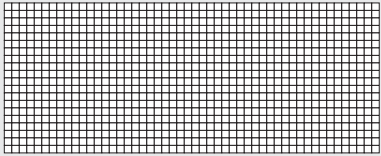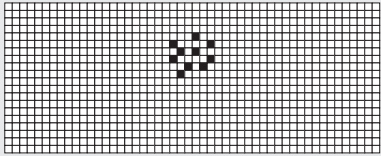
Figure 1 Example of stated-preference trade-off question.

perforated bowel. MAR is defined as the maximum probability of experiencing a treatment-related AE that subjects are, on average, willing to accept to obtain a given level of symptom improvement.

Specifying all attribute levels as discrete variables, the MAR for AE *i* for an improvement in symptoms can be defined as:

$$\sum_{j} S_j^1 - \sum_{j} S_j^0 = -\left(\sum_{k=1}^{m_j} R_k^i + f_{ij} \times R_{m_j+1}^i\right) \quad 0 \le f_{ij} < 1$$

$$\text{MAR}^i(1,0) = r_{m_j}^i + f_{ij} \times r_{m_j+1}^i$$

where S_j^0 is the estimated preference parameter for the initial set of symptom levels, and S_j^1 is the estimated preference parameter for the improved set of symptom levels. We calculate the amount of risk that exactly offsets the benefit of improved symptoms. R_k^i is the preference parameter for risk type *i*. If the offsetting value lies between $R_{m_j}^i$ and $R_{m_j+1}^i$, we solve for fraction f_{ij} to linearly interpolate between the two discrete risk levels. Finally, we calculate MAR as $r_{m_j}^i + f_{ij} \times r_{m_j+1}^i$, where r_m^i is the AE probability corresponding to preference parameter R_m^i. This calculation is directly analogous to estimating willingness to pay with a non-linear cost function.

Results

Survey Population

Of all the individuals who were eligible, 589 individuals completed the online questionnaire. Of these 589 individuals, 13 did not vary their choices across the trade-off questions (i.e., they picked Medicine A or Medicine B in every choice question).

Exclusion of these subjects resulted in an analysis sample of 576 patients. Table 2 summarizes the personal characteristics and IBS experience of these subjects. The demographic characteristics of the subjects were similar to those of the women enrolled in two trials of alosetron efficacy and safety [20,21].

Preference Weights

Preference weights for the attribute levels are presented in Figure 2 (The effects-coded, random-parameters estimates used to derive preference weights are available from the authors). Estimates are naturally ordered except for no AE risk and 0.1% chance of moderate colitis. This difference was statistically insignificant. The treatment attribute with the highest relative importance weight was perforated-bowel risk, followed by severe-colitis risk, frequency of abdominal pain, and diarrhea, respectively. The least important attribute was the risk of moderate colitis. The relative importances of the two most important attributes were not statistically significantly different from each other, but these two attributes were significantly more important than the other attributes ($P < 0.05$).

MAR

MARs for the three serious AEs for each of the four defined levels of benefit are presented in Table 3 (MARs for other combinations of symptom outcomes are available from the authors by request). MAR estimates ranged from 0.03% to 2.83%. As expected, subjects in our sample were willing to accept higher levels of risk in return for greater improvements in symptoms. For each level of treatment benefit, MAR estimates were highest for impacted bowel and lowest for perforated bowel, indicating that subjects

Table 2 Patient demographic characteristics and irritable bowel syndrome (IBS) experience

Characteristic	Sample (N = 576)
Age, mean (SD), year	47 (12)
Race/ethnicity, n (%)	
White	546 (95)
Black/African American	6 (1)
Native American or Alaska Native	7 (1)
Other	17 (3)
Employment status, n (%)	
Employed for wages full time	170 (30)
Employed for wages part-time	76 (13)
Self-employed	33 (6)
Homemaker	89 (15)
Student	13 (2)
Retired	71 (12)
With disability/unable to work	100 (17)
Unemployed, but looking for work	24 (4)
Highest level of education, n (%)	
High school or equivalent	80 (14)
Trade school	27 (5)
Some college but no degree	222 (39)
Associate's degree	92 (16)
Bachelor's/college degree or higher	155 (27)
How long since diagnosed, mean (SD), year	10 (9)
Severity of IBS symptoms	
Mild	26 (5)
Mild to moderate	117 (20)
Moderate	186 (32)
Moderate to severe	214 (37)
Severe	33 (6)
Symptoms experience with IBS, n (%)	
Frequent and severe abdominal pain and discomfort	462 (80)
Frequent feeling of urgency	441 (77)
Limitations on daily activities	365 (63)

had greater tolerance for the relatively less serious risk of impacted bowel than for the much more serious risk of perforated bowel.

We compared the risk tolerance of subjects who reported serious IBS symptoms (serious pain or discomfort, urgency or soiling of clothes, or IBS symptoms that impact their daily activities) or who took or had taken prescription medications to treat IBS with subjects who did not have these characteristics. Patients with more serious symptoms have higher MARs than patients who had less serious symptoms. For example, for full benefit, subjects with more serious symptoms are willing to accept 3.76% and 1.57% risks of impacted and perforated bowel, respectively. In contrast, subjects with less serious symptoms were willing to accept only 1.64% and 1.19% risks ($P < 0.01$ for both comparisons).

Discussion

As expected, women's choices indicated a systematic preference for treatments that provide larger reductions in symptom frequency. In many instances, preferences for symptom relief outweighed concerns about AE risks. For example, the estimated preference weight for no abdominal pain and discomfort is greater than the estimated preference weight for no risk of serious AEs, indicating that relief of abdominal pain and discomfort was more important to these women than is eliminating the AE risks.

Variation in the MAR measure of risk tolerance was consistent with the relative seriousness of the AEs. For example, for a treatment offering moderate symptom improvement relative to severe symptoms, MAR for impacted bowel was 1.53%, whereas

MAR for perforated bowel was only 0.86%. The estimated MARs for clinically meaningful symptom improvements are substantially above their estimated rates of occurrence across each of the serious AEs of interest [22].

Most importantly, estimating preference weights for treatment outcomes and risks, and quantifying the relative importance of each attribute offer a solution to the problem in benefit–risk analysis that risks and benefits are measured in non-comparable units. In a companion article, these preference weights are used in the first incremental net benefit analysis to use such patient data to compare treatment benefits and risks [5].

One inherent limitation of this methodology is that SC questions ask subjects to evaluate hypothetical treatments. Thus, differences can arise between stated and actual choices. In the present study, potential hypothetical bias is minimized by offering alternates that mimic real-world trade-offs as closely as possible.

Subjects enrolled through the Ipsos Observer were not screened to confirm their reported IBS diagnosis. We consider it is unlikely that people who do not have IBS completed the survey because no personal gain was associated with participation in the survey, other than a chance to win one of the five $100 cash incentives. In addition, participation required a commitment of approximately 30 minutes to complete the survey, and the SC trade-off tasks are mentally taxing exercises that are unlikely to attract a casual respondent.

Health status is based on patients' own report of a physician diagnosis of IBS. We have no independent verification of that diagnosis. Nevertheless, we followed up the self-report with questions on type and severity of symptoms—abdominal pain and discomfort, feeling of urgency or soiling clothes, and inability to lead a normal home or work life because of the need to be near a bathroom. About 80% of the subjects indicated one or more of these symptoms were "frequent," and 75% judged the severity to be moderate or severe.

As indicated, this study provided weights for a separate modeling study to estimate incremental net benefits. We thus were constrained to obtain weights that matched the end points in clinical-trial data. Rates for each outcome are reported, and thus modeled, independently. For example, subjects sometimes evaluated outcome profiles that included both IBS-related diarrhea and medication-induced constipation in the same week. The effect of a given diarrhea frequency might be different, depending on the combination of diarrhea days and constipation days in that week. This interaction effect is not reflected in our importance estimates for these end points. Conventional health-state utility weights widely used in cost-effectiveness analyses also treat outcomes as independent and are subject to the same limitation. One could argue that models and importance weights should account for interactions among outcomes, but they do not because the data on which models are based generally do not provide the necessary information to model such interactions. An important topic for future research would be to relax the independence requirement in both the model and the preference weights to evaluate the significance of this requirement.

Patients' perspectives on balancing benefits and risks may be useful in informing both treatment and regulatory decisions. Because risks, often, are inseparable from efficacy, one cannot easily define what level of risk is intolerable without reference to the benefits associated with increased risk. SC studies such as this one may help decision-makers understand the levels of risks that patients are willing to accept in return for therapeutic benefits. Quantitative estimates of preferences for combinations of risks and benefits developed using rigorous and theoretically sound techniques, such as those used in this study, may assist regulatory
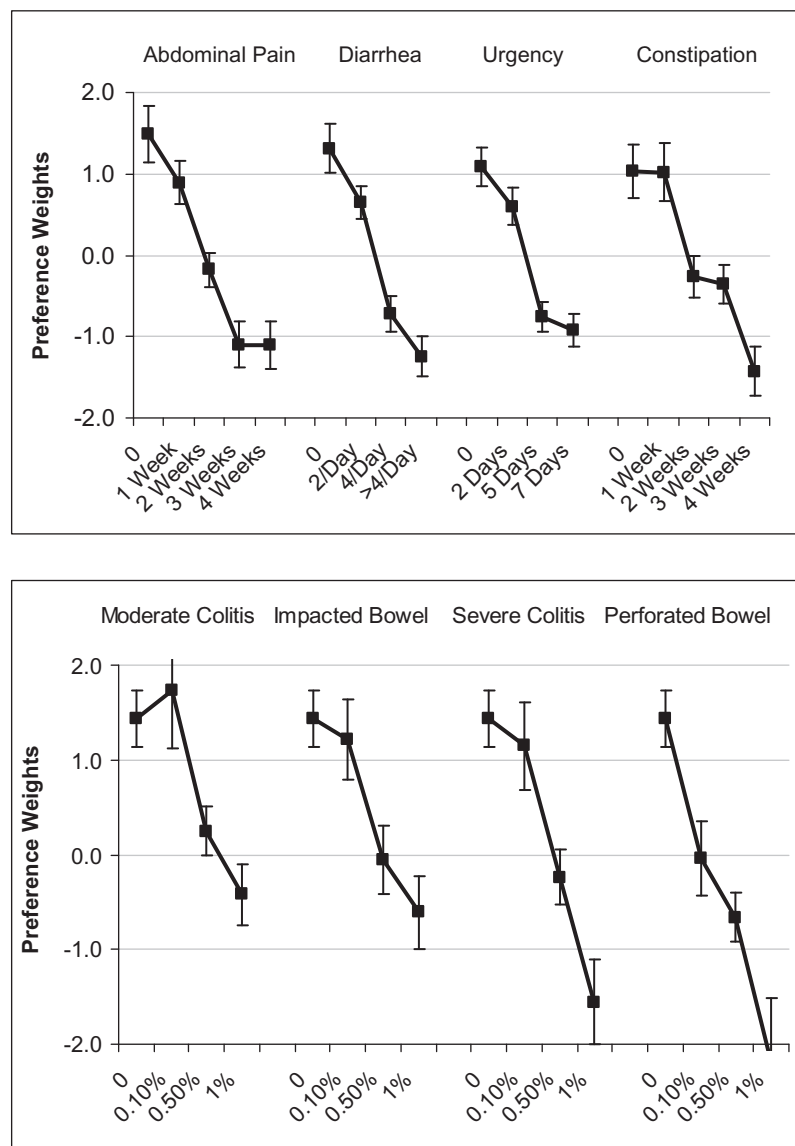
Figure 2 Preference weights.

Table 3 Maximum acceptable risk (MAR) for each serious adverse event for different levels of treatment benefit

Symptom improvement*	MAR % (95% confidence interval) for each type of risk		
	Impacted bowel	Severe colitis	Perforated bowel
Complete symptom relief	2.83 (1.96, 6.93)	2.01 (1.51, 3.46)	1.37 (1.22, 1.57)
Good improvement	2.03 (1.48, 4.50)	1.43 (1.13, 2.34)	1.10 (1.01, 1.21)
Moderate improvement	1.53 (1.15, 3.15)	1.07 (0.83, 1.60)	0.86 (0.68, 1.03)
Poor improvement	0.17 (0.02, 0.38)	0.14 (0.01, 0.35)	0.03 (0.00, 0.07)

*MARs are estimated for symptom improvements relative to the worst combination of abdominal pain, diarrhea, and urgency. Severe symptoms were defined as the combination of the worst levels of abdominal pain, diarrhea, and urgency. Poor symptom improvement was defined as abdominal pain for 3 weeks, diarrhea 4 times each day, and urgency 5 days each week. Moderate symptom improvement was defined as abdominal pain for 2 weeks each month, diarrhea 2 times each day, and urgency 2 days each week. Good symptom improvement was assumed to be the same as moderate symptom improvement with regard to diarrhea and urgency, but included only 1 week of abdominal pain each month. Complete symptom relief was defined as no abdominal pain, diarrhea, urgency, or constipation.

authorities in evaluating new treatments, making the rationale for decisions more transparent and helping physicians and patients make better-informed choices among treatments.

References

1 Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002: the importance of reporting suspected reactions. Arch Intern Med 2005; 165:1363–9.
2 Hirst C, Cook S, Dai W, et al. A call for international harmonization in therapeutic risk management. Pharmacoepidemiol Drug Saf 2006;15:839–49.
3 Institute of Medicine. The Future of Drug Safety: Promoting and Protecting the Health of the Public. Washington, DC: Institute of Medicine (U.S.), 2007.

4 Wolfe SG, Chey WY, Washington MK, et al. Tolerability and safety of alosetron during long-term administration in female and male irritable bowel syndrome patients. Am J Gastroenterol 2001;96:803–11.

5 Lynd LD, Najafzadeh M, Colley L, et al. Using the incremental net benefit framework for quantitative benefit–risk analysis in regulatory decision-making—a case study of alosetron in irritable bowel syndrome. Value Health 10 September 2009 [Epub ahead of print].

6 Johnson FR, Ozdemir S, Manjunath R, et al. Factors that affect adherence to bipolar disorder treatments—a stated-preference approach. Med Care 2007;45:545–52.

7 Ryan M, Netten A, Skatun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. J Health Econ 2006;25:927–44.

8 Ashcroft DM, Seston E, Griffiths CEM. Trade-offs between the benefits and risks of drug treatment for psoriasis: a discrete choice experiment with UK dermatologists. Br J Dermatol 2006;155:1236–41.

9 Johnson FR, Özdemir S, Mansfield C, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. Gastroenterology 2009;133:769–79.

10 Johnson FR, Ozdemir S, Hauber B, Kauf TL. Women's willingness to accept perceived risks for vasomotor symptom relief. J Womens Health 2007;16:1028–40.

11 Kuhfeld WF, Garratt M, Tobias RD. Efficient experimental design with marketing research applications. J Mark Res 1994;31:545–57.

12 Dey A. Orthogonal Fractional Factorial Designs. New York: Halstead Press, 1985.

13 Huber J, Zwerina KB. The importance of utility balance in efficient choice designs. J Mark Res 1996;33:307–17.

14 Kanninen B. Optimal design for multinomial choice experiments. Journal of Marketing Research 2002;39:214–27.

15 Zwerina K, Huber J, Kuhfeld WF. A general method for constructing efficient choice designs. SAS Working Paper, 2006.

16 Ipsos Observer. Data Collection. Available from: http://www.ipsos-na.com/observer/ [Accessed June 2008].

17 Train K. Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press, 2003.

18 Gauss 7.0 [Computer Program]. Black Diamond, WA: Aptech Systems, Inc., 2005.

19 Hensher DA, Rose JM, Greene WH. Applied Choice Analysis: A Primer. Cambridge: Cambridge University Press, 2005.

20 Camilleri M, Chey WY, Mayer EA, et al. A randomized controlled clinical trial of the serotonin type 3 receptor antagonist alosetron in women with diarrhea-predominant irritable bowel syndrome. Arch Intern Med 2001;161:1733–40.

21 Camilleri M, Northcutt AR, Kong S, et al. Efficacy and safety of alosetron in women with irritable bowel syndrome: a randomised, placebo-controlled trial. Lancet 2000;355:1035–40.

22 Mangel AW, Northcutt AR. Review article: the safety and efficacy of alosetron, a 5-HT$_3$ receptor antagonist, in female irritable bowel syndrome patients. Aliment Pharmacol Ther 1999;13 (Suppl. 2):77–82.