B. Statistical Methods

1. Respondent Universe and Sampling Methods

Participants will be consumers who self-identify as having been diagnosed with one of three possible medical conditions: depression, high cholesterol, or insomnia. All participants will be 18 years of age or older. We will exclude individuals who work in healthcare or marketing settings because their knowledge and experiences may not reflect those of the average consumer. Recruitment and administration of the study will take place over the Internet. Participation is estimated to take approximately 30 minutes.

Research Now's opt-in online survey panel is demographically balanced, including racial and ethnic minorities, a wide range of different age groups, and individuals with relatively less educational attainment. They recruit panel members through a combination of e-mail, online marketing, and by invitation, with over 300 diverse online and offline affiliate partners and targeted website advertising. By using multiple recruitment methods Research Now is able to recruit a diverse set of representative consumers and decision makers to participate in their panels. Panel inclusion is by invitation only, and Research Now invites only prevalidated individuals with known characteristics to participate in the consumer panels.

Recently, Research Now has launched a "Mental Health Panel," providing access to a targeted audience of over 260,000 panelists in the US who identify themselves as suffering from depression, epilepsy, bipolar disorder, narcolepsy, sleep disorder or sleep apnea, and anxiety. Although the database is always growing and changing, the current demographic distribution of the database is presented in Exhibit 1. Using Research Now's database and a screening questionnaire, we will recruit 2,100 individuals for the pretest and main study combined.

Exhibit 1. Current Demographic Distribution of Respondent Database

Demographic Characteristic	Percent
Gender	
Female	59%
Male	41%
Age	
18–24	10%
25–34	24%
35–44	21%
45–54	19%
55–64	17%
65 or over	9%

FDA does not intend to generate nationally or locally representative results or precise estimates of population parameters from this study. The sample used is a convenience sample, rather than a probability sample. Despite the attempt to match between the study's sample and known population characteristics, matching is used solely to produce samples with a reasonable degree of diversity in key demographic characteristics. Furthermore, no legitimate weights can be constructed from non-probability samples such as the one used here. Hence, the Agency does not construe this sample or the results generated from this sample as nationally or locally representative. Rather, the strength of the experimental study lies in its internal validity, on which meaningful estimates of differences across conditions can be produced and generalized.

2. Procedures for Collection of Information

All parts of this study will be administered over the Internet. Within medical condition, participants will be randomly assigned to view one of four possible versions of an ad, as depicted in Table 1 below. One version will present the full major statement without the disclosure regarding additional risks (Conditions C, G, and K). This version will implement existing ads in

the marketplace. Stimuli variations for the other three versions will be achieved by replacing the audio track of the original ad with the revised risk and disclosure statements described above. Thus, a second version of the ad will include the full major statement plus the disclosure about additional risks (Conditions A, E, and I). A third version will include an abbreviated statement of risks without the disclosure about additional risks (Conditions D, H, and L). The fourth version will include an abbreviated statement of risks as well as the disclosure about additional risks (Conditions B, F, and J).

After viewing the ad, participants will respond to questions about information in the ad. Measures are designed to assess perception and understanding of product risks and benefits; perception and understanding of the disclosure about additional risks; perceptions of product quality; intention to seek more information about the product; and perceptions of attitudes toward advertising disclosures. Demographic and health literacy information will be collected. In addition, participants will answer questions about their familiarity with their medical condition. The entire procedure is expected to last approximately 30 minutes in the pretest and main study. This will be a one-time (rather than annual) information collection. The questionnaire is included in Appendix 2.

Table 1: Study Design

		Major Statement	
Medical Condition	**Disclosure Regarding Additional Risks**	**Version 1**	**Version 2**
Depression	Present	A	B
	Absent	C	D
High Cholesterol	Present	E	F
	Absent	G	H
Insomnia	Present	I	J
	Absent	K	L

Note. Version 1 = current major statement; Version 2 = abbreviated major statement

Analysis Plan

Our primary hypothesis for this study is that, relative to inclusion of the full major statement, providing limited risk information along with the disclosure about additional risks will promote improved consumer perception and understanding of serious and actionable drug risks. We will also investigate other questions such as whether overall drug risk and benefit perceptions are affected by these changes. Other exploratory analyses will be conducted as well.

We will conduct ANOVAs (for continuous variables) and chi-squares and logistic regressions (for categorical variables) to examine the impact of abbreviated risk statements and disclosures. Before conducting analyses, we will assess whether the inclusion of covariates is justified. If they are, we will conduct the analyses both with and without covariates (e.g., sex, age, race/ethnicity, education) included in the model. If the interaction effect (risk statement x disclosure) is significant, we will conduct pairwise-comparisons to determine which conditions

are significantly different from one another. The primary planned comparison will use a p-value of .05, and the exploratory post-hoc comparison will use Bonferroni-adjusted *p*-values.

Power

The following assumptions were made in deriving the sample size for the main studies and the pretests: (1) 0.90 power, (2) 0.05 alpha level, and (3) an effect size between moderately small and medium. For the main studies, our power analysis suggests that, for each medical population, a sample size of 125 per experimental group will allow for the required 90% power (alpha = 0.05 using two-sided testing) to assess the effects of the four stimuli (representing two types of manipulations). Specifically, results from a sensitivity analysis for a 2x2 ANOVA suggests that this design can detect moderately small effects (f =.15).[1] An effect size of f = 0.15 is traditionally considered moderately small, whereas an effect size of f=0.25 is considered medium.[2] If we were to alter this analysis to include four covariates (such as gender, age, race/ethnicity, education), the resulting two-way ANCOVA would be slightly less sensitive. It would still be capable of detecting conventionally moderately small effects (f = .16). We also conducted a sensitivity analysis for the primary planned comparison (comparing the abbreviated risk statement / disclosure group to the full risk statement / no disclosure group). Assuming .90 power, an alpha level of .05, and equal sample sizes in each comparison group (n = 125), independent samples t tests will be able to detect "moderately small" sized effects (d = .41). Finally, we conducted a sensitivity analysis for the post-hoc comparisons between the remaining experimental arms. Assuming .90 power, an alpha level of .01 (Bonferroni-adjusted for five comparisons), and equal sample sizes in each comparison group (n = 125), independent samples t tests will detect medium-sized effects (d = .49).

1 In a 2x2 ANOVA, the numerator degrees of freedom for the two main effects and the interaction term are all equal to 1. Thus, the overall sample size required to detect each of these effects is the same.

2 Jacob Cohen (1988), *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), New Jersey: Lawrence Erlbaum Associates.

For analyses involving discrete outcome variables (e.g., correctly understood risks vs. incorrectly understood risks), the proposed sample size will allow detection of an absolute difference of 21 percentage points in arm-to-arm comparisons (e.g., a difference between 0.71% in one arm versus 0.50% in the comparison arm) with a power of 0.90. In this calculation, we assumed equal-sized samples in each arm (n = 125), a design effect equal to 1, an alpha level of 0.05, a two-sided Fisher's exact test, and an underlying proportion of study participants in a particular response category equal to 0.50. An underlying proportion of 0.50 is the most conservative estimate and overestimates the sample size relative to alternate proportions.

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

Response rates can vary greatly depending on many factors including the sample composition, panel type, invitation content, time of day and incentive offering. In addition, outside factors including email filters, recipient ISP downtime and general conditions on the Internet can impact response rates. We will calculate response rate as ratio of the number of surveys completed to the number of panelists contacted by invitation. To help ensure that the participation rate for the internet panel is as high as possible, FDA and the contractor will:

- Design an experimental protocol that minimizes burden (short in length, clearly written, and with appealing graphics);
- Administer the experiment over the Internet, allowing respondents to answer questions at a time and location of their choosing;
- Send out two email reminders after the initial invitation;
- Provide respondents with a helpdesk link that they can access at any time for assistance.

Additionally, the Panel leverages the social media concept and has developed 'panel communities' in order to maximize member engagement and overcome challenge of declining survey response rates and multi-panel membership. We will also conduct a demographic comparison of responders and non-responders and incorporate any findings into our discussion of results.

4. Test of Procedures or Methods to be Undertaken

Two types of pretesting (qualitative and quantitative) will be employed as a test of procedures and methods.[3] The first type of pretesting will be qualitative. Cognitive testing with nine individuals will be used to refine study questions. Following cognitive testing, one round of quantitative pretesting will be employed. Pretests will be used to refine the design of the experimental stimuli to ensure the validity of the manipulations. These pretests will explore variations in the design of the stimuli, such as lines, color, and layout. The main study design will not change as a result of the pretests, nor will the results of the pretests increase the burden on respondents in the main study. The pretests are designed to ensure the particulars of the main study are implemented in the best way possible.

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

The contractor, RTI International, will collect and analyze data on behalf of FDA as a task order under Contract HHSF223201400272G. Vanessa Boudewyns, Ph.D., is the Project Director, 202-728-2092 (x22092). Review of contractor deliverables and supplemental analyses will be provided by the Research Team, Office of Prescription Drug Promotion (OPDP), Office

3 Pretesting is suggested by OMB as a method to test procedures. See Office of Management and Budget *Standards and Guidelines for Statistical Surveys* (September, 2006). Available at http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf. Last accessed January 12, 2012.

of Medical Policy, CDER, FDA, and coordinated by Kevin R. Betts, Ph.D., WO BLDG 51, RM 3220, (240) 402-5090, and Kathryn J. Aikin, Ph.D., WO BLDG 51, RM 3240, (301) 796-1200.