B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. **Respondent Universe and Sampling Methods**

The universe for this experimental study is members of the Knowledge Networks Internet panel. Knowledge Network's Internet panel consists of 48,000 adult panel members who are systematically recruited by random-digit dialing (RDD) or by using address-based sampling. Households without existing Internet service are also eligible, and Knowledge Networks provides these members with laptops or Web TVs to enable their participation. The sample is nationally representative and statistically accurate. Typically, panel members receive 3-4 invitations per month to participate in research projects.

For the study, 4,650 (900 for the pretests, 3,750 for the main study) participants will be recruited from the panel. Each panel member will complete a prescreening questionnaire (see Appendix 2 for the screener and Appendix 4 for the recruitment and reminder emails). The target population is the adult noninstitutionalized population in the US with access to the internet, are age 18 and over, do not work for a pharmaceutical company, an advertising agency, a market research company, or are healthcare professionals. The contractor will recruit participants with the goal of yielding 250 participants in each of the 15 test conditions for a total of 3,750 completed interviews. Upon entry into the survey, participants will be screened for age and occupation. The goal will be to have a final sample with similar race/ethnicity and education proportions to the US population, though we will not be attempting to make generalizable population point estimates from these data. The interpretation of results will include a discussion of the generalizability of the findings given this sampling method.

The resulting sample is representative of the target population by definition due to the probabilistic multistage process for the construction of the panel and stratified random selection of panel members. Due to propensities to respond and other random processes, the sample may not be reflective of the population, i.e., the sample distribution may be slightly skewed relative to the population. This will be corrected using statistical weights. The weights will be constructed such that the weighted sample distribution matches population as described by the most recent Current Population Survey for key demographic categories.

We are not limiting recruitment to individuals with specific medical conditions. This should avoid problems with familiarity of details of the medical condition. Should we end up with individuals who possess the medical condition in the sample, we plan to use medical condition as a covariate.

After qualifying for the survey, each respondent will be randomly assigned to an experimental condition. Assignment to condition only after qualifying for the survey ensures equal and unbiased allocation of the respondents to experimental condition.

2. **Procedures for Collection of Information**

Design Overview

We will investigate the effects of adding disease outcome information to branded promotional materials on consumer perceptions and understanding. This information will be examined in the context of direct-to-consumer (DTC) prescription drug print advertisements. We hope to more readily generalize our findings by exploring the issues raised above in three medical conditions varying in severity and symptomatology: COPD, lymphoma and anemia.

We plan to examine two variables in this study: the type of disease information (possible disease outcomes, versus non-outcome information, versus no information) and the format of the information (integrated with drug information versus separated). Some participants will see information about the disease that avoids discussion of disease outcomes the drug has not been shown to address, such as, "Diabetes is a disease in which blood sugar can vary uncontrollably, leading to uncomfortable episodes of high or low blood sugar." Other participants will see disease information that mentions consequences of the disease that go beyond the indication of the advertised product, such as, "Untreated diabetes can lead to blindness, amputation, and, in some cases, death." A third group will see drug product information only (no disease information). We will also examine the way in which the disease information is presented relative to the product claims in the piece by varying the format: disease information mixed (integrated) with product claims versus disease information apart (separated) from product claims. We are exploring a number of different options for implementing these two variables. For example: alternating paragraphs of product and disease information, disease information on one page and product information on another page, use of different colors and fonts for disease and product information, and different visuals for disease and product information. Final format variations will be determined through pretesting. The pretests are designed only to make sure the particulars of the main study are implemented in the best way possible. The results of the pretests will not increase the burden on respondents in the main study, nor will the main study design change as a result of the pretests.

This study utilizes random assignment to conditions. Within medical condition, participants will be randomly assigned to see one version of the ad. Participants will be

recruited from a general population sample to control for prior knowledge about disease outcomes.

The design is described in Table 1:

Table 1.--- Study Design

		Format of Disease and Product Information		
Medical Condition	Disease Information plus	Integrated	Separated	Control (no disease info)
COPD	Non-outcome	1	2	5
	Outcomes	3	4	
Lymphoma	Non-outcome	6	7	10
	Outcomes	8	9	
Anemia	Non-outcome	11	12	15
	Outcomes	13	14	

Procedure

This study will be administered over the internet. A total of 4,650 interviews will be completed. Participants will be randomly assigned to view one version of a DTC prescription drug print ad which consists of a one- or two-page display section and the accompanying brief summary page. Following their perusal of this document, they will answer questions about their recall and understanding of the benefit and risk information, their perceptions of the benefits and risks of the drug, and their intent to ask a doctor about the medication.

Demographic and subjective health literacy information[1] will be collected. The entire procedure is expected to last approximately 20 minutes. This will be a one-time (rather than annual) information collection.

Participants

For the study, 4,650 (900 for the pretests, 3,750 for the main study) participants will be recruited from an online panel. Each panel member will complete a prescreening questionnaire (see Appendix 2 for the screener and Appendix 4 for the recruitment and reminder emails). Because the task presumes basic reading abilities, all selected participants must speak and read English fluently. Participants must be 18 years or older.

Hypotheses

1. Individuals presented with disease outcome information will be more likely to inaccurately recall (misattribute) outcome information as a product benefit compared to individuals who see non-outcome information or the control ad. This effect will be strongest when disease outcome information is integrated with product information. We do not expect any differences on risk recall.

2. Individuals presented with integrated disease outcome information will have the highest perceived product efficacy, perceive the risk/benefit balance to be tilted more toward benefit, and have the greatest behavioral intentions compared to individuals who

1 Objective measures of health literacy are typically long (e.g., S-TOFHLA) or require an interviewer (e.g., REALM). The item we used to measure subjective health literacy is based on previous work (Chew et al., 2004) and has been validated (Jeppesen et al., 2009; Morris et al., 2006). Chew, L.D., Bradley, K.A., Boyko, E.J. (2004). Brief questions to identify patients with inadequate health literacy. *Family Medicine, 36*, 588-594; Jeppesen, K.M., Coyle, J.D., Miser, W.F. (2009). Screening questions to predict limited health literacy: A cross-sectional study of patients with diabetes mellitus. *Annals of Family Medicine, 7*, 24-31; Morris, N.S., MacLean, C.D., Chew, L.D., & Litternberg, B. (2006). The Single Item Literacy Screener: Evaluation of a brief instrument to identify limited reading ability. *BMC Family Practice, 7*, 21-27.

see non-outcome information or the control ad. This effect will be strongest when disease outcome information is integrated with product information.

3. Individuals presented with integrated disease outcome information will have the lowest perceived risk compared to individuals who see non-outcome information or the control ad; that is, as perceived efficacy increases, perceived risk will decrease because, according to the Affect Heuristic (Slovic & Peters, 2006),[2] people perceive things that are more beneficial as less risky.

4. Individuals presented with disease outcome information will be more likely to get the product's indication wrong compared to individuals who see non-outcome information or the control ad. This effect will be strongest when disease outcome is integrated with product information.

5. Individuals presented with disease outcome information will have more confidence in their judgment compared to individuals who see non-outcome information or the control ad. This effect will be strongest when disease outcome is integrated with product information.

6. We will explore the effects tested in the hypotheses to see if they are modified by demographic variables. For instance, individuals who have the medical condition may respond differently than individual who do not have the condition. All other analyses are exploratory.

Analysis Plan

We will conduct ANOVAs (for continuous variables) and chi-squares and logistic regressions (for categorical variables) to test the hypotheses outline above. We will

2 Slovic, P. and E.Peters, (2006). Risk Perception and Affect. *Current Directions in Psychological Science, 15*, 322-325.

conduct these analyses both with and without covariates (e.g., demographic and health characteristics) included in the model. If a main effect is significant, we will conduct pairwise-comparisons to determine which conditions are significantly different from one another, with Bonferroni-adjusted *p*-values.

Power Analysis

We have calculated the sample size for the pretests and study using power analysis. The following table shows the power calculations for the study. The assumptions made in deriving the sample size for the study were: 1) 0.90 power, 2) 0.05 alpha and a 0.005 Bonferroni-adjusted alpha, and 3) an effect size between small and medium. The tables below shows the sample size required to detect differences with effect sizes ranging from conventionally "small" (f = 0.10) to "medium" (f = 0.25).

Table 8: Power Analysis Calculation: Main Study

A priori power analysis to determine sample size needed in F tests (ANOVA: fixed effects, main effects, and interactions) to achieve power of 0.90 (Faul et al., 2007).[3]							
		Main effect of information Effect size f*			Post-hoc comparisons among cells Effect size f*		
Input							
		0.10	0.15	0.25	0.10	0.15	0.25
	α error probability	0.05	0.05	0.05	0.005**	0.005**	0.005**
	Power (1 – β error probability)	0.90	0.90	0.90	0.90	0.90	0.90
	Numerator df	1	1	1	1	1	1
	Number of groups	4	4	4	2	2	2
Output							
	Critical F	3.85	3.86	3.90	7.90	7.93	8.01
	Denominator df	1049	465	167	1671	742	267
	Sample size per	**263**	**117**	**43**	**838**	**374**	**136**

3 Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A, (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

	cell						

*An effect size of 0.10 is traditionally considered small, whereas an effect size of 0.25 is considered medium (Cohen, 1988).[4] Here we have shown three different effect sizes centered around small to medium effects.

**Bonferroni-adjusted for 10 comparisons.

In the main study, we will have 250 participants per cell, for a total of 3,750 participants in the study (a 2 x 2 + 1 design for three different medical conditions). This will provide sufficient power to detect small to medium effects for main effects and multiple comparisons.

The following assumptions were made in deriving the sample size for the pretests: 1) 0.90 power, 2) 0.10 alpha and a 0.02 Bonferroni-adjusted alpha, and 3) an effect size between small and medium. Table 9 shows the sample size required to detect differences with three different effect sizes ranging from f = 0.10 to f = 0.25.

Table 9: Power Analysis Calculation: Pretests

A priori power analysis to determine sample size needed in F tests (ANOVA: fixed effects, main effects, and interactions) to achieve power of 0.90							
		Main effect Effect size f*			Post-hoc comparisons among cells Effect size f*		
Input							
		0.10	0.15	0.25	0.10	0.15	0.25
	α error probability	0.10	0.10	0.10	0.02**	0.02**	0.02**
	Power (1 – β error probability)	0.90	0.90	0.90	0.90	0.90	0.90
	Numerator df	3	3	3	1	1	1
	Number of groups	4	4	4	2	2	2
Output							
	Critical F	2.09	2.09	2.11	5.42	5.44	5.49

4 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed)*. Hillsdale, NJ: Lawrence Erlbaum & Associates, Inc.

	Denominator df	1179	524	188	1303	580	210
	Sample size per cell	**296**	**132**	**48**	**653**	**291**	**106**

*An effect size of 0.10 is traditionally considered small, whereas an effect size of 0.25 is considered medium (Cohen, 1988).[5] Here we have shown three different effect sizes centered around small to medium effects.

**Bonferroni-adjusted for 10 comparisons.

We propose to examine four stimuli variations to refine the stimuli for each medical condition. With 75 participants per cell, we could detect medium effects for the main effect and large effects for comparisons, for a total of 900 participants in the pretest (four variations in three medical conditions).

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;
- Sending out two email reminders after the initial invitation (see Appendix 4).

5 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed)*. Hillsdale, NJ: Lawrence Erlbaum & Associates, Inc.

- 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.

This procedure has been reviewed and approved by FDA's human subject protection committee (RIHSC).

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.[6] The first type of pretesting will be qualitative. Cognitive testing with nine individuals will be used to refine study questions. Following cognitive testing, two rounds 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.

---

6 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.

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

The contractor, Synovate, Inc., will collect the information on behalf of FDA as a task order under Contract HHSF223200910136G. Valerie Fuller-DiPaula, Ph.D., is the Project Director for this project, 703-663-7243. Data analysis will be conducted primarily by the Research Team, Office of Prescription Drug Promotion (OPDP), Office of Medical Policy, CDER, FDA, and coordinated by Kathryn J. Aikin, Ph.D., (COTR), WO BLDG 51, RM 3240, (301) 796-1200 and Helen W. Sullivan, Ph.D., M.P.H., WO BLDG 51, RM 3252, (301) 796-1200.