B. Statistical Methods (used for collection of information employing statistical methods)

1. Respondent Universe and Sampling Methods

The eligible study population is U.S., non-institutionalized adults age 18 and older. The selected sample will be drawn from GfK Custom Research's Internet panel (KnowledgePanel) according to the specific research objectives for this project. The KnowledgePanel consists of approximately 55,000 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 GfK provides these members with laptops and Internet access to enable their participation.

The survey sample will be drawn from eligible members using an implicitly stratified systematic sample design based on the methodology for which GfK was assigned a U.S. Patent (U.S. Patent No. 7,269,570) in September 2007. The selection methodology, which has been used by GfK since 2000, assures that GfK panel samples will closely track the U.S. population, and that survey panelists will not be over-burdened with survey requests. Typically, panel members receive 2-3 invitations per month to participate in research projects.

The sample will be drawn from KnowledgePanel members who report suffering from chronic pain or high blood pressure for the first two phases and the general population for the second two phases of iterative testing and will be approximately evenly divided between men and women.

In the unlikely event that insufficient KnowledgePanel members will be available to create a sample of adequate size for chronic pain and hypertension, GfK will obtain additional sample from a reputable off-panel vendor (e.g., Research Now).

Eligible participants for the pretest and the medical condition sample of the main study will be adults who speak English and self-identify as having been diagnosed with chronic pain or high blood pressure (hypertension). Eligible participants for the general population sample of the main study will be adults who speak English. We will exclude individuals who work in the health care, marketing, advertising, or pharmaceutical industries. We will also exclude pretest study participants from the main study.

2. Procedures for the Collection of Information

Design Overview

We will conduct pretesting prior to main data collection to assess the psychometric properties and identify any measurement challenges (e.g., misinterpretation, lack of variance) with candidate measurement items. We also will use the pretesting to examine factors that may affect future study results and analyses (e.g., response scale midpoints, moderating variables). We will conduct two sequential pretest waves (n=500 per wave; n=1,000 total) with the following target populations: (a) individuals diagnosed with chronic pain and (b) individuals diagnosed with hypertension.

Exhibit 1. Pretest Study Design

Wave	Medical Condition		
	Chronic Pain	Hypertension	
Wave 1	n = 250	n = 250	**500**
Wave 2	n = 250	n = 250	**500**
TOTAL	**500**	**500**	**1,000**

In the main study phase, we will conduct four sequential waves of iterative testing to fully assess the measurement properties of the candidate items and create the final pool of measurements. We will conduct the first two waves of the main study with members of the target populations (hypertension and chronic pain) to refine the measurement items for those groups and the second two waves with members of the general population who do not have the target health conditions to determine if measurement reliability and validity change when the advertised drug addresses a condition that study participants do not have (n=2,500 per wave; n=10,000).

Exhibit 2. Iterative Testing Design – Illness Population Sample

Wave 1									
Chronic Pain Ad					**Hypertension Ad**				
Ad Type	**Drug Risk Level**	**Drug Benefit Level**		**Control**	**Ad Type**	**Drug Risk Level**	**Drug Benefit Level**		**Control**
		High	**Low**				**High**	**Low**	
Print	**High**	n=125	n=125	n=125	**Print**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	
Television	**High**	n=125	n=125	n=125	**Television**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	
Wave 2									
Chronic Pain Ad					**Hypertension Ad**				
Ad Type	**Drug Risk Level**	**Drug Benefit Level**		**Control**	**Ad Type**	**Drug Risk Level**	**Drug Benefit Level**		**Control**
		High	**Low**				**High**	**Low**	
Print	**High**	n=125	n=125	n=125	**Print**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	
Television	**High**	n=125	n=125	n=125	**Television**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	

Exhibit 3. Iterative Testing Design – General Population Sample

Wave 3									
Chronic Pain Ad					**Hypertension Ad**				
Ad Type	**Drug Risk Level**	**Drug Benefit Level**		**Control**	**Ad Type**	**Drug Risk Level**	**Drug Benefit Level**		**Control**
		High	**Low**				**High**	**Low**	

Print	**High**	n=125	n=125	n=125	**Print**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	
Television	**High**	n=125	n=125	n=125	**Television**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	

Wave 4

Chronic Pain Ad					**Hypertension Ad**				
Ad Type	**Drug Risk Level**	**Drug Benefit Level**		**Control**	**Ad Type**	**Drug Risk Level**	**Drug Benefit Level**		**Control**
		High	**Low**				**High**	**Low**	
Print	**High**	n=125	n=125	n=125	**Print**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	
Television	**High**	n=125	n=125	n=125	**Television**	**High**	n=125	n=125	n=125
	Low	n=125	n=125			**Low**	n=125	n=125	

Procedure

Pretests: Each participant will be randomly assigned to view either a print ad or a television ad for a fictitious prescription drug indicated to treat chronic pain or hypertension and will be asked to complete a brief online survey assessing their benefit/risk perceptions, intentions, and attitudes toward the drug. Based on the pretest findings, we will revise and remove candidate items prior to full-scale testing.

Main study: Each participant will be randomly assigned to view either a print or television ad for a fictitious prescription drug for hypertension or chronic pain and will be asked to complete a brief online survey assessing their benefit/risk perceptions, intentions, and attitudes toward the drug. In the first two main study waves, participants will view an ad that matches the sample's medical condition (chronic pain or hypertension). In the final two main study waves, participants will be randomly assigned to view either the chronic pain stimuli or the high blood pressure stimuli.

Participants

Eligible participants for the pretest and the medical condition sample of the main study will be adults who speak English and self-identify as having been diagnosed with chronic pain or high blood pressure (hypertension). Eligible participants for the general population sample of the main study will be adults who speak English. We will exclude individuals who work in the health care, marketing, advertising, or pharmaceutical industries. We will also exclude pretest study participants from the main study.

Analysis Plan

Overall: We will conduct a series of analyses on each wave of the iterative testing to assess the psychometric properties of the measurement items, select items for the final measure pool, develop and validate scales, and construct appropriate scoring algorithms.

For each population (general and illness-specific), our analysis process will include both a development (exploratory) stage to investigate the properties of the items and develop scales and a validation stage to confirm the findings in the development stage and assess the validity of the scales.

Wave 1: We will examine the percentage of participants endorsing each response option —as well as the number of missing, don't know, or not applicable responses—to identify items that may be confusing, have little variability in responses, or demonstrate floor or ceiling effects. Next, we will conduct exploratory factor analyses to investigate the factor structure of the items and determine how the items might be grouped into scales. The most appropriate number of factors will be selected based on eigenvalues, percentage of variance accounted for by the factors, demonstration of simple structure in the factor pattern (i.e., factors load highly on only one factor), and interpretability of the factors.

After establishing the factor structure of the items, we will conduct item response theory (IRT) analyses to further examine properties of the items and scales. After identifying the best-fitting IRT model, we will conduct IRT-based differential item functioning (DIF) analyses to determine if item properties are consistent across the subgroups of interest (e.g., gender, racial/ethnic, education groups). At the end of Wave 1, we will retain the best-performing items for further testing.

Wave 2: In addition to repeating the IRT and DIF analyses, we will conduct confirmatory factor analyses (CFA) to test whether the items cluster into factors as expected based on the Wave 1 findings. If the factor models fit well, we will conduct multi-group confirmatory factor analyses to assess the invariance of the factor structure across the subgroups defined by mode, gender, race/ethnicity, and education. If the model does not fit well, we will examine the residuals and modification indices to identify sources of misfit and perform exploratory factor analyses to determine an appropriate factor structure. Based on the factor analysis results, we will combine the items into scales and will develop appropriate scoring algorithms.

We will then assess the reliability and validity of scales. We will compute Cronbach's alphas to assess the internal consistency reliability of each scale for both the overall sample and separately for the subgroups. To assess the construct validity of the scales, correlations of the newly constructed risk and benefit scales with the other previously validated measures will be computed. The new scales should correlate most highly with scales measuring similar constructs (convergent validity) and should have lower correlations with scales measuring different constructs, such as perceived worry (discriminant validity). Results demonstrating this pattern provide support for the validity of the scales. To further examine the validity of the scales, we will conduct analyses of covariance to compare mean scale scores of participants in different experimental conditions, after controlling for demographic factors (e.g., race/ethnicity, gender, age, education). For example, mean risk scale scores should be significantly higher for participants in the high risk conditions than those in the low risk conditions after controlling for other factors.

Wave 3: In Wave 3, we will repeat similar analyses (descriptive, IRT, DIF) as conducted in Wave 1 for the general population. However, given the prior work in Waves 1 and 2 establishing the factor structure of the scales, we will conduct confirmatory, rather than exploratory, factor analyses to directly test whether the factor structures found in the illness-specific populations also apply to the general population.

Wave 4: We will conduct similar analyses to Wave 3. In addition to verifying the properties of the items in this wave using IRT and factor analyses, we will assess the reliability and validity of the scales in the general population. Cronbach's alphas will be computed to assess internal consistency of each scale. Construct validity will be evaluated based on correlations between the new scales and other scales of similar and dissimilar constructs and analyses of covariance to compare scale scores across experimental conditions, controlling for other factors.

Power

As shown in Exhibits 2 and 3 above, each wave of the main study will include 2,500 participants (125 per cell). We estimated power for analyses of variance (ANOVAs) comparing continuous outcome measures by medical condition (chronic pain, hypertension), ad type (print, television), drug risk level (high, low), and drug benefit level (high, low, control) at p-value of 0.05, using the PASS software program (Hintze, 2011)[1] and testing the following effect sizes for ANOVA specified by Cohen (1988)[2]: small (f=0.10), medium (f=0.25), and large (f=0.40). With the projected sample size, we would have greater than 90% power for small, medium, and large-sized main and interaction effects.

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

This experimental study will use an existing research panel to draw a sample. The panel comprises individuals who have signed up to participate in online studies. To help ensure that the participation rate is as high as possible, FDA will:

- Design an experimental protocol that minimizes burden (clearly written and with appealing graphics);
- Administer the survey over the Internet, allowing respondents to answer questions at a time and location of their choosing;
- Provide e-mail reminders throughout the course of the field period;
- Make telephone reminder calls to nonrespondents;
- Provide a toll-free hotline for respondents who may have questions or technical difficulty as they complete the survey.

There are several approaches to address the potential for nonresponse bias analysis in this study, such as comparing response rates by subgroups, comparing respondents and nonrespondents on frame variables, and conducting a nonresponse follow-up study.[3] For

[1] Hintze, J. (2011). PASS 11. NCSS, LLC. Kaysville, Utah, USA. www.ncss.com.
[2] Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
[3] Office of Management and Budget, *Standards and Guidelines for Statistical Surveys*, September, 2006.

the proposed project, we will compare responders and nonresponders on demographic variables.

4. Test of Procedures or Methods to be Undertaken

   In a previous data collection (OMB Control Number 0910-0497; "Risk and Benefit Perception Scale Development Focus Groups") we conducted focus groups to explore how individuals think about prescription drug risks and benefits, and what influences their decision to take (or not take) a particular drug. The results of these groups were used to develop measures that, in addition to measures identified through our literature search, were subsequently tested in cognitive interviews (OMB Control Number 0910-0695; "Cognitive Interviews for Risk and Benefit Perception Scale Development Study"). The cognitive testing examined the draft measures to refine and improve question wording and narrowed the pool of questions. In addition to these steps, we will conduct pretesting to test and further refine the measurement pool to be used in the main study.

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

   The contractor, RTI, will collect and analyze the data on behalf of FDA as a task order under Contract HHSF223201210375G. Bridget Kelly, Ph.D., M.P.H., 202-728-2098, is the Project Director for this project. Data analysis will be overseen by the Research Team, Office of Prescription Drug Promotion (OPDP), Office of Medical Policy, CDER, FDA, with assistance from the Office of Biostatistics, CDER, and coordinated by Kathryn J. Aikin, Ph.D., 301-795-0569, and Helen W. Sullivan, Ph.D., M.P.H., 301-796-4188.

www.whitehouse.gov/sites/default/files/omb/inforeg/statpc. Last accessed April 18, 2013.