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

1. Respondent Universe and Sampling Methods

The selected sample will be drawn from a respondent panel managed by Toluna, a digital market research and technology company that offers targeted access to well-profiled, highly engaged physicians and allied healthcare professionals for global research needs. Toluna's Healthcare Practice maintains proprietary panels of physicians, nurses, pharmacists, and allied healthcare professional in the United States and the EU5 (France, Germany, Italy, Spain, United Kingdom) who have opted in to receiving invitations to participate in market research.

Currently, Toluna has access to 5,500 oncologists and 11,950 primary care physicians (PCPs). We will use Toluna's healthcare panel called Curizon. This is a dedicated respondent community used exclusively for healthcare research. The Curizon panel will be supplemented with partner panels, an email and direct mail campaign to the American Medical Association (AMA) list of healthcare professionals to reach the total required sample size.

The Toluna Affiliate Network is used to augment their proprietary panels. Toluna only works with respondent providers who have their own proprietary panels. This means they typically work with local suppliers who are based in the country under study. Suppliers follow all industry guidelines; provide adequate responses to ESOMAR's 28 Questions; and partners are asked for opt-in policies, privacy policies, legal guidelines, and geographic locations. Once a supplier is onboard, it is automatically and continuously monitored to gauge performance.

For this study, Toluna has designated the following primary partners: WebMD and M3 Global Research. Toluna anticipates being able to obtain approximately 500 oncologists from the three panels combined (as well as the total PCP and advanced practice practitioner [APP] populations). To supplement the remainder of the sample, Toluna is proposing a direct mail and email campaign by renting an AMA list. This list will have access to both email and postal addresses, totaling 15,707 postal addresses and 11,713 email addresses.

2. Procedures for the Collection of Information

Part A of the supporting statement described the rationale for conducting the study.

The proposed study seeks to address the following research questions:

1. Do disclosures mitigate potentially misleading presentations of preliminary and/or descriptive data in oncology drug product promotion?

2. Does the language (technical, non-technical) of the disclosure influence the effectiveness of the disclosure?

3. Does the presence of a general statement about the clinical utility of the data in addition to a specific disclosure influence processing of claims and disclosures?

4. Do primary care physicians and oncologists differ in their processing of claims and disclosures about preliminary and/or descriptive data? Does oncology experience alter processing of claims and disclosures about preliminary or descriptive data?

5. Which disclosures do physicians prefer?

To address these questions, FDA has designed a study that will be conducted in three independent phases, each phase examining a data display in a promotional piece for a unique oncological product. Independent variables will include a) specific disclosure (technical, non-technical, none), b) general statement (present, absent), and c) specialty (oncologists, primary care physicians). Each phase will have the following design:

Specialty	**General Statement**	**Specific Disclosure**		
		Technical	**Non-technical**	**No Disclosure**
Oncologists	Present	•	•	Control
	Absent	•	•	
Primary Care Physicians	Present	•	•	Control
	Absent	•	•	

Specific disclosures will include material information specifically related to the particular data display in question. As such, each specific disclosure may include clinical or statistical information related to the trial design, the statistical analysis plan of the trial, or any other material statistical or clinical information necessary for evaluation or interpretation of the data. The team developing the disclosures includes social science analysts, pharmacists, oncology medical officers, and an oncology nurse. An example of the general statement is “This presentation includes exploratory information of uncertain clinical utility and should be interpreted cautiously when used to make treatment decisions.”

Outcome variables will focus on the assessment of the data display as a whole as well as attention to the disclosure, if present. Specifically, we will examine recognition of the clinical endpoint in the data display, comprehension of the data display, perceptions of the exploratory nature of the data, and the perceived credibility of the promotional piece. We will also look at attention to the specific disclosure and the general statement, prescriber decisions, and prescriber preferences. This latter outcome variable will be determined by a secondary task at the end of the questionnaire that shows each participant all disclosure options and asks them to choose their preferred version.

Oncologists and primary care physicians will be recruited to participate via Internet and the study is expected to take approximately 20 minutes. Participants will view professionally developed promotional pieces that mimic currently available promotion and answer questions.

Specific Hypotheses and Research Questions (by Dependent Variable)

1. Recognition of clinical endpoint in data display

- **RQ1**: Does specialty and oncology experience affect recognition of the clinical endpoint in the data display?

- **RQ2**: Will experimental condition (five levels: control vs. general-present plus technical vs. general-absent plus technical vs. general-present plus nontechnical vs. general-absent plus nontechnical) affect recognition of the clinical endpoint in the data display?

2. Comprehension of data display

- **HYP1**: Participants exposed to a specific disclosure (technical or nontechnical) will have greater data display comprehension than participants not exposed to a specific disclosure (control group).

- **HYP2**: Participants exposed to a nontechnical disclosure will have greater data display comprehension than participants exposed to a technical disclosure.

- **HYP2b**: Oncologists exposed to a technical disclosure will have greater data display comprehension whereas PCPs will have greater data display comprehension when exposed to non-technical disclosure (interaction effect).

- **HYP3**: Participants exposed to a general statement and a specific disclosure will have greater data display comprehension than participants exposed to a specific disclosure without a general statement.

- **HYP4**: Specialists will have greater data display comprehension than PCPs.

3. Perceived exploratory nature of data

- **HYP5**: Participants exposed to a general disclosure will have greater agreement that the data are exploratory or uncertain than participants not exposed to a general disclosure.

- **RQ3**: Does experimental condition (five levels: control vs. general-present plus technical vs. general-absent plus technical vs. general-present plus nontechnical vs. general-absent plus nontechnical) affect perceptions of the exploratory/uncertain nature of the data?

- **RQ4**: Does specialty and oncology experience affect perceptions of the exploratory/uncertain nature of the data?

4. Perceived credibility of promotional piece

- **HYP6**: Participants exposed to a technical disclosure will perceive the stimuli as more credible and scientific than participants exposed to a nontechnical disclosure.
- **RQ5**: Will sales aids with no general or specific disclosures (control group) have the lowest perceived credibility compared to all other groups combined?

5. Attention to specific disclosure

- **HYP7**: Participants exposed to a specific disclosure (either technical or nontechnical) will be more likely to recall seeing a specific disclosure than those not exposed to a specific disclosure (control group).

- **HYP8**: Participants exposed to a nontechnical disclosure will have greater specific-disclosure recall than participants exposed to the technical disclosure.

- **HYP9**: Participants exposed to a general statement and a specific disclosure will have greater recognition of specific disclosures than participants exposed to a specific disclosure without a general statement.

- **RQ6**: Will specialty and oncology experience affect whether participants notice the specific disclosure?

6. Attention to general statement

- **HYP10**: Participants exposed to a general statement will be more likely to recall seeing a general statement than those not exposed to a general statement.

- **RQ7**: Will specialty and oncology experience affect whether participants notice the general statement?

7. Benefit perceptions

- **HYP11**: Participants in the control condition will perceive the drug as more beneficial than participants exposed to a general statement or a specific disclosure.

- **RQ8**: Does specialty and oncology experience affect benefit perceptions?

8. Prescribing decisions

- **HYP12**: Participants not exposed to specific or general disclosures (control group) will report greater intentions to prescribe the advertised drug than those exposed to disclosures (all other groups combined).

- **RQ9**: Will the presence of a general statement affect intentions to prescribe?

9. Disclosure preferences

- **RQ10**: When asked whether a data display should include either a general statement or a specific disclosure, which type of disclosure will participants prefer (and why)?

10. Time spent viewing stimuli

- **RQ11**: How long will providers take to read the promotional piece (e.g., time spent viewing promotional piece)? Will this differ by specialty and oncology experience?

Analysis Plan

Descriptive Analysis

During descriptive analysis, we will calculate frequency distributions and check the apparent validity of the data (i.e., range checks, frequency of missing responses, or response distribution). For continuous/ordinal variables, statistical output will include means, medians, standard deviations, ranges, and counts. For categorical variables, output will include counts and percentages.

In addition to frequency distributions, we will conduct three other types of analyses during this step. First, we will calculate reliability of composite variables and multi-item scales to determine if the individual items hang together as composite measures. Specifically, we will calculate Cronbach's alpha for each composite variable. If alpha for a composite measure or scale does not meet our pre-established threshold of 0.75, we will discuss whether to use single-item measures rather than the composite or to consider such composites as indices (because of a theoretical reason to consider an aggregate measure regardless of item correspondence) in hypothesis testing.

Finally, we will conduct a non-response analysis to compare the distribution patterns of responders with known population distributions. These comparisons will be limited to those variables for which we have population information (e.g., sex, age, race/ethnicity).

Hypothesis Testing

We will test hypothesized relationships implied by our central research questions by conducting one of several statistical tests as outlined below. We will examine associations between demographic and background characteristics with outcomes of interests through chi-square tests for categorical outcomes and t-tests for continuous outcomes to control within subsequent models for such characteristics with notable differences across groups. For analyses involving the control group, we will treat the 2×2×2+1 design as a 2 (specialty: oncologist, PCP) × 5 (general statement and specific disclosure: technical disclosure/statement absent, technical disclosure/statement present, nontechnical disclosure/statement absent, nontechnical disclosure/statement present, control). For each study phase, we will implement two-way ANOVAs to test for significant differences in continuous outcome variables among experimental groups. The two-way ANOVA models—and multinomial models for categorical outcomes—will determine whether there is a main effect of training level or the five-level disclosure factor. In the main study, we will also explicitly test for simple effects by specific disclosure (technical, non-technical, none) and general statement (present, absent).

We will define a set of planned contrasts to address specific research questions and hypotheses, conducting these planned comparisons based on hypothesized relationships or post-hoc comparisons to identify significant differences between specific experimental groups. To adjust for multiple comparisons, we will apply a post-hoc family-wise error-rate adjustment, such as a Bonferroni correction. If assumptions are violated for the ANOVA or categorical models mentioned above, nonparametric tests will be employed, such as the Kruskal-Wallis for independent samples or Welch's ANOVA. Both alternatives help detect statistical differences should assumptions such as normality or equal variances not be met.

In the main test, statistical output will include, as appropriate, F or chi-square statistics, test degrees of freedom, p values, mean or proportional differences, and standardized effect sizes.

Power

Pretest

The proposed sample size for the pretest is 60 (n = 30 of each specialty). No power analysis was conducted because the purpose of the pretest will be to test the survey administration process and assess survey timing, as shown in **Table 2**.

Table 2. Pretest Sample Size by Phase and Specialty

	PCPs	Oncologists	Total
Phase 1: Promotional Piece #1	*n* = 10	*n* = 10	***N* = 20**
Phase 2: Promotional Piece #2	*n* = 10	*n* = 10	***N* = 20**
Phase 3: Promotional Piece #3	*n* = 10	*n* = 10	***N* = 20**
Total			***N* = 60**

Note. PCP = Primary care physician

Main Studies

We have powered the main studies to detect moderately small effects. For analyses involving the control group, we will treat the 2×2×2+1 design as a 2 (specialty: oncologist, PCP) × 5 (general statement and specific disclosure: technical disclosure/statement absent, technical disclosure/statement present, nontechnical disclosure/statement absent, nontechnical disclosure/statement present, control). Assuming power of 0.90 and an alpha = 0.05, our omnibus tests will be able to detect an effect size of $f = 0.12$ for the main effect of the specialty group and $f = .15$ for the main effect of the five-level disclosure factor. Given these assumptions, we will need a total of $N = 705$ for each study phase (see **Table 3**).

Table 3. Main Study Target Sample Size per Condition (*N* = 705) (to be repeated for each study phase)

	General disclosure statement		No Disclosure	Specific Disclosure	
				Technical	NonTechnical
Specialty	**Oncologist**	Absent	*n* = 47	*n* = 47	*n* = 47
		Present		*n* = 47	*n* = 47
	PCP	Absent	*n* = 94	*n* = 94	*n* = 94
		Present		*n* = 94	*n* = 94

Note. PCP = Primary care physician

With the same parameters, we will also be able to conduct planned contrasts testing various combinations of group means with enough sensitivity to detect moderately small effects ($f = 0.12$) for orthogonal contrasts, where no family-wise error correction is required because each contrast is statistically independent. The design is also sensitive to detect moderately small size effects ($f = 0.16$) for up to 14 nonorthogonal contrasts, assuming a Bonferroni-adjusted alpha of 0.0035. This study is also powered to detect pairwise interactions for the two manipulated factors (disclosure and specialty) in a nested analysis excluding the control group. For categorical outcomes, this sample will give us 90% power to detect a small effect size of $w = 0.15$ using a 4 degree-of-freedom chi-square test with an alpha significance level of 0.05.

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

Both the pretest and main studies will be administered via Internet. To help ensure that the participation rate is as high as possible, FDA and the contractor will:

- Design a protocol that minimizes burden (short in length, clearly written, and with appealing graphics);

- Use incentive rates that meet industry standards. In addition to offsetting respondent burden, using market-rate incentives tends to increase response rates, reduce sampling bias, and reduce nonresponse bias.

- Use government sponsorship on the survey invite to increase response rate. An experiment conducted by FDA and RTI[1] found that among endocrinologists, response rates were 6 percentage points higher when FDA was disclosed as the sponsor in the survey invitation than when no sponsor was listed. However, due to concerns raised in the public comments that mentioning FDA could potentially influence subjects' responses to study questions, we will ensure that all materials reference the U.S. Department of Health and Human Services rather than FDA.

Participants in the pretest and main studies will be convenience samples, rather than probability-based samples of U.S. adults. Rather, the strength of the experimental design used in this study lies in its internal validity, on which meaningful estimates of differences across manipulated conditions can be produced and generalized. This is a counterpoint to observational survey methodologies where estimating population parameters is the primary focus of statistical analysis. The recruitment procedures in this study are not intended to fit the criteria for survey sampling, where each unit in the sampling frame has an equal probability of being selected to participate. In an observational survey study, response rates are often used as a proxy measure for survey quality, with lower response rates indicating poorer quality. Nonresponse bias analysis is also commonly used to determine the potential for nonresponse sampling error in survey estimates. However, concerns about sampling error do not generally apply to experimental designs, where the parameters of interest are under the control of the researcher—rather than being pre-established characteristics of the participants—and each participant has an equal probability of being assigned to any of the experimental conditions.

Generally, there are several approaches to conducting a nonresponse bias analysis, such as comparing response rates by subgroups, comparing respondents and nonrespondents on frame variables, and conducting a nonresponse follow-up study[2]. For the proposed project, we will examine nonresponse for its descriptive value by comparing our full sample with population estimates for age, race, gender.

4. Test of Procedures or Methods to be Undertaken

[1] Aikin, KJ; Betts, K; Boudewyns, V; Stine, A; & Southwell, B. (2016). Physician responsiveness to survey incentives and sponsorship in prescription drug advertising research. *Annals of Behavioral Medicine, 50*(Suppl.), s251.

[2] Office of Management and Budget, *Standards and Guidelines for Statistical Surveys*, September, 2006. www.whitehouse.gov/sites/default/files/omb/inforeg/statpc. Last accessed April 18, 2013.

Nine cognitive interviews were conducted to assess questionnaire flow and wording. We plan to conduct a pretest on a larger scale to ensure the main studies will run smoothly. We propose to test 90 individuals in the pretest.

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 HHSF223201510002B. Vanessa Boudewyns, Ph.D., 202-728-2092, 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, and coordinated by Amie C. O'Donoghue, Ph.D., 301-796-0574, and Kathryn J. Aikin, Ph.D., 301-796-0569.