Tobacco Products: Principles for Designing and Conducting Tobacco Product Perception and Intention Studies

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments regarding this draft guidance may be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Electronic comments may be submitted to https://www.regulations.gov. Alternatively, submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Room 1061, Rockville, MD, 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft guidance, contact the Center for Tobacco Products at (Tel) 1-877-CTP-1373 (1-877-287-1373) Monday-Friday, 9 a.m. – 4 p.m. EDT.

Additional copies are available online at http://www.fda.gov/TobaccoProducts/Labeling/RulesRegulationsGuidance/default.htm. You may send an e-mail request to SmallBiz.Tobacco@fda.hhs.gov to receive an electronic copy of this guidance. You may send a request for hard copies to U.S. Food and Drug Administration, Center for Tobacco Products, Attn: Office of Small Business Assistance, Document Control Center, Bldg. 71, Rm. G335, 10903 New Hampshire Ave., Silver Spring, MD 20993-2000.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Tobacco Products

October 2020

Table of Contents

Tobacco Products: Principles for Designing and Conducting Tobacco Product Perception and Intention Studies

Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

This draft guidance is intended to help applicants (or "you") design and conduct tobacco product perception and intention (TPPI) studies that may be submitted as part of a modified risk tobacco product application (MRTPA), a premarket tobacco product application (PMTA), or a substantial equivalence report (SE Report). TPPI studies are studies that can be used to assess, among other things, individuals' perceptions of tobacco products, understanding of tobacco product information (e.g., labeling, modified risk information), and intentions to use tobacco products. It is possible for a TPPI study to also include an actual use component (e.g., an actual product utilized in a simulated use setting or a real environment of use); however, a discussion of actual use research is beyond the scope of this draft guidance. This draft guidance addresses the following scientific issues for applicants to consider as they design and conduct TPPI studies to support tobacco product applications:

- Developing TPPI study aims and hypotheses
- Designing quantitative and qualitative TPPI studies
- Selecting and adapting measures of TPPI study constructs
- Determining TPPI study outcomes

[1] This draft guidance was prepared by the Office of Science and Office of Regulations in the Center for Tobacco Products at FDA.

- Selecting and justifying TPPI study samples
- Analyzing TPPI study results

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidance documents describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word should in Agency guidance documents means that something is suggested or recommended, but not required. This guidance provides non-binding recommendations on TPPI studies and does not establish requirements for submitting studies in support of an application.

II. BACKGROUND

The Federal Food, Drug, and Cosmetic Act ("FD&C Act") generally requires new tobacco products to undergo review and receive authorization from FDA before being introduced or delivered for introduction into interstate commerce. The FD&C Act establishes three premarket review pathways for new tobacco products:

- Submission of a PMTA under section 910(b) and receipt of a marketing order under section 910(c)(1)(A)(i);
- Submission of a SE report under section 905(j)(1)(A) and receipt of an SE marketing order; and
- Submission of a request for an exemption from SE under section 905(j)(3) and receipt of an exemption from FDA (implemented at 21 CFR § 1107.1).

In addition, under section 911(a) of the FD&C Act, prior to marketing a modified risk tobacco product (MRTP), an applicant must submit a modified risk tobacco product application (MRTPA) and receive an order under section 911(g) of the FD&C Act.

The results of TPPI studies can help an applicant demonstrate that its new tobacco product meets the applicable premarket authorization standard.[2] In addition, the results of TPPI studies can help an applicant demonstrate that its proposed modified risk tobacco product[3] satisfies the

[2] For more information about the PMTA pathway and how TPPI studies may be used in these types of applications to help demonstrate that the marketing of the tobacco product is appropriate for the protection of public health, please see the guidance document "Premarket Tobacco Product Applications for Electronic Nicotine Delivery Systems" at https://www.fda.gov/tobacco-products/rules-regulations-and-guidance/guidance. While the guidance is written about submitting PMTAs for a specific product type, it provides some general information about the PMTA pathway that may be useful for all PMTA applicants. For more information about the substantial equivalence pathway and how TPPI studies may help applicants demonstrate that a new tobacco product is substantially equivalent to a predicate product, please see the guidance documents "Section 905(j) Reports: Demonstrating Substantial Equivalence for Tobacco Products" and "Demonstrating the Substantial Equivalence of a New Tobacco Product: Responses to Frequently Asked Questions (Revised)" at https://www.fda.gov/tobacco-products/rules-regulations-and-guidance/guidance. In addition, FDA has promulgated a final rule on exemptions from substantial equivalence requirements, codified at 21 CFR 1107.1.

[3] FDA has issued a draft guidance document "Modified Risk Tobacco Product Applications" which provides information about the types of scientific studies and analyses FDA recommends that applicants consider conducting

standard in section 911(g) of the FD&C Act. The results of these studies on individuals' perceptions of tobacco products, understanding of information about tobacco products, or intentions to use tobacco products may help predict future tobacco use behavior. For example, an SE Report may include a TPPI study that demonstrates the potential effect of a difference in pouch size between the new and predicate smokeless products on reported intentions for tobacco use behaviors in different populations. The outcomes of TPPI studies can also provide information about the likelihood that nonusers will initiate use with the new product, and likelihood that users will change their tobacco use behavior ((e.g., use an MRTP in addition to, or in place of, their current tobacco product). TPPI studies can also be used to evaluate whether people understand the label, labeling, and advertising of a product, and whether it is misleading. For example, TPPI studies can be used in MRTPAs to help demonstrate the product's labeling or advertising enables the public to understand the modified risk information and the relative significance of the information in the context of total health.[4]

FDA encourages applicants considering development of TPPI studies to request a meeting with FDA to discuss their research and development plans related to their tobacco products.[5]

III. DEFINITIONS

For the purposes of this guidance document, FDA intends to use the following definitions:

Comparison product means the product or products (including product categories) to which the applicant seeks to compare the tobacco product that is the subject of the application.[6] For example, in an MRTPA requesting a claim that the proposed modified risk tobacco product is less harmful than another commercially marketed product, the other commercially marketed product would be a comparison or comparator product.

Label is defined in section 201(k) of the FD&C Act, which states in part that the term label means a display of written, printed, or graphic matter upon the immediate container of any article.

Labeling is defined in section 201(m) of the FD&C Act and means all labels and other written, printed, or graphic matter (1) upon any article or any of its containers or wrappers, or (2) accompanying such article.

---

to provide the evidence needed to support the issuance of an order under section 911(g) of the FD&C Act. This draft guidance, when finalized, will represent FDA's current thinking on this topic. It is available at https://www.fda.gov/tobacco-products/rules-regulations-and-guidance/guidance.

[4] Section 911(h)(1) of the FD&C Act.

[5] For additional information on meetings with FDA, please see the Guidance for Industry and Investigators, "Meetings with Industry and Investigators on the Research and Development of Tobacco Products" at https://www.fda.gov/tobacco-products/rules-regulations-and-guidance/guidance.

[6] This term is inclusive of predicate tobacco products (*i.e.*, a tobacco product commercially marketed (other than for test marketing) in the United States as of February 15, 2007, or a tobacco product that FDA has previously determined to be substantially equivalent to which the applicant claims its new tobacco product is substantially equivalent).

Measures, in the context of a TPPI study, are questions or items that study participants respond to and that are intended to serve as indicators of a particular construct that is being investigated within the study.

Modified risk tobacco product (MRTP) means any tobacco product that is sold or distributed for use to reduce the harm or the risk of tobacco-related disease associated with commercially marketed tobacco products (section 911(b)(1) of the FD&C Act).

New Tobacco Product means any tobacco product (including those products in test markets) that was not commercially marketed in the United States as of February 15, 2007; or any modification (including a change in design, any component, any part, or any constituent, including a smoke constituent, or in the content, delivery, or form of nicotine, or any other additive or ingredient) of a tobacco product where the modified product was commercially marketed in the United States after February 15, 2007 (section 910(a)(1) of the FD&C Act).

Perception is an umbrella term used regularly in public health research for a cluster of related, but distinct, psychological constructs, including: beliefs, attitudes, judgments, and expectancies.

Stimuli, in the context of a TPPI study, includes the materials (e.g., label, labeling, or advertising) presented to study participants, about which the participants are asked to respond. This includes actual (physical) products, images of products, product labels, labeling, advertising, or other marketing materials.

Surveys are instruments for gathering data by systematically asking study participants questions and recording their answers.

Tobacco Product is defined in section 201(rr) of the FD&C Act, which states in part:

> (1) The term "tobacco product" means any product made or derived from tobacco that is intended for human consumption, including any component, part, or accessory of a tobacco product (except for raw materials other than tobacco used in manufacturing a component, part, or accessory of a tobacco product).
>
> (2) The term "tobacco product" does not mean an article that is a drug under section 201(g)(1) of the FD&C Act, a device under section 201(h) of the FD&C Act, or a combination product described in section 503(g) of the FD&C Act.

IV. OVERALL APPROACH

When developing TPPI studies, there are important general principles to consider regarding study design and method, study personnel, and the relationship between the product(s) included in the study and the product(s) you plan to make the subject of an application.

First, FDA recommends you employ the design and method that will most effectively achieve your TPPI study aims. TPPI studies can use various combinations of designs (e.g., experimental,

longitudinal, cross sectional) and methods (qualitative (e.g., focus groups) and quantitative (e.g., self-administered surveys)). For example, TPPI studies could be conducted using experimental surveys, or cross-sectional, structured interviews.

We recommend applicants conduct studies using best practices specific to the study design and method employed that have been either (1) written or published by well-established social or behavioral scientific organizations (e.g., the American Psychological Association), or (2) written by authors who have demonstrated scientific expertise in the type of method or design (i.e., have published studies using the method or design in peer-reviewed journals).[7] Adhering to best practices improves the validity of the data collected and the conclusions drawn from those data. For example, it is a best practice that experimental studies use random assignment to assign participants to control and experimental conditions. Random assignment improves the study's internal validity by ruling out alternative explanations for any differences detected between conditions.

Second, we recommend that applicants select appropriate study personnel for TPPI studies. Appropriate study personnel may include study personnel involved in the design, implementation, and analysis of the TPPI study who have sufficient formal education, training, and experience in conducting social or behavioral science research to ensure the study is designed and conducted appropriately. Study personnel should be able to recognize and address features of a TPPI study that could introduce bias or compromise the validity of the study results (e.g., poorly worded measures, failure to "blind" study personnel and participants to condition assignment or study hypotheses). Using appropriate study personnel helps ensure that the data and analyses are scientifically valid. This includes considerations related to the disclosure of conflicts of interest (e.g., disclosure of financial ties and the role of the study sponsor in the design and conduct of the research) and decisions about "blinding" (e.g., blinding personnel or participants to condition assignment or study hypotheses). Moreover, adherence to ethical principles for conducting research that involves human subjects, acceptable to the research and public health communities, should be followed to ensure that adequate procedures are in place for protection of human subjects. Adequate procedures for human subject protection help protect the rights, safety, and welfare of human subjects and ensure the integrity and scientific validity of the study data.

Third, if you plan to use the results of the study as support for an application for more than one product, you should include each product in the study when feasible. When this is not feasible (e.g., you have many varieties of a product), you should select products to include based on a rationale that describes why results based on the tested products can be generalized to products that were not tested.

[7] *See e.g.*, Rosenthal, R., & Rosnow, R. L. (1991). Essentials of Behavioral research: Methods and Data Analysis (Vol. 2). New York: McGraw-Hill.

V. DEVELOPING STUDY AIMS AND HYPOTHESES

FDA recommends that you develop TPPI study aims (the overall goals of the study) prior to conducting the study. These aims should consider how your study informs your application and how the study relates to the standard for authorization. For example, an applicant who believes marketing their tobacco product as an MRTP may benefit the population as a whole could conduct a TPPI study to assess, among other things, whether smokers would likely start using the proposed modified risk product and stop smoking cigarettes, and whether nonsmokers would likely not initiate use of the product. One aim of the TPPI study could be to assess the effect of exposure to the labeling and advertising for the proposed modified risk product on intentions to try the product among smokers and nonsmokers. Results could help inform FDA's assessment of the potential impact of the modified risk tobacco product on use behavior in tobacco users and nonusers, which is relevant to determining whether the product meets the standard for authorization.

We also recommend that you develop specific hypotheses (or research questions, if you do not have hypotheses) that can be tested statistically by analyzing study data. These hypotheses provide information on how the study data will be used to address the study aims. We recommend you classify hypotheses as primary and secondary, which will inform the power analysis and sample size estimation (see Section XI.C. of this guidance).

In some studies, you might hypothesize that there will be no difference between groups or no relationship between variables. For instance, for a PMTA, you may hypothesize that there will be no differences in intentions to use a new product between persons who plan to quit using tobacco products and those who do not. In such cases, it is especially important to have documented that the measures chosen for the study are valid measures of the constructs being investigated, and that the study is sufficiently statistically powered to detect differences should they exist (see Sections IX and XI.C of this guidance for discussions of measurement and power, respectively). This helps FDA rule out two common alternative explanations for null findings—poor quality measures and low statistical power.

VI. QUALITATIVE STUDY METHODS

Qualitative study methods include, for example, focus group and in-depth interview studies. In general, FDA recommends you consider using both qualitative and quantitative methods when conducting TPPI research. While a single quantitative study may provide adequate support for your application, there are advantages to conducting multiple studies using various methods. For example, conducting qualitative studies can be useful to develop modified risk information that can be later tested in quantitative studies.[8]

A. Qualitative Method Considerations

[8] Qualitative and quantitative methods can also be combined for a mixed-methods approach. *See e.g.,* National Institutes of Health, Office of Behavioral and Social Science Research, *Mixed Methods Research*, https://obssr.od.nih.gov/training/online-training-resources/mixed-methods-research/.

As previously mentioned, examples of qualitative methods include focus groups[9] and in-depth interviews.[10]

When conducting a qualitative study, FDA recommends the following:

- Developing a guide, in advance of conducting the study, that includes all the questions and potential probes to be used during the interviews or focus groups to ensure consistency in the procedures across moderators or interviewers and across focus groups or interviews. The format and structure of the guide may vary depending on the goal of study; for instance, some guides include specific probes while others allow for a more open-ended, reactive discussion. Regardless of the format, a guide helps to reduce bias that could be introduced by the moderator or interviewer because it standardizes the language used;
- Recording and transcribing each focus group or interview for analysis, with participants' consent;
- Employing additional strategies to minimize the possibility that moderators or interviewers will introduce bias:
    - The moderators or interviewers of focus groups and interviews should be appropriately trained in qualitative research techniques. For example, trained moderators/interviewers avoid encouraging or discouraging a particular response by using techniques such as responding to participant comments using neutral language and facial expressions.
    - When feasible, moderators or interviewers should be blind to study hypotheses, if there are any, to prevent them from affecting the data collected (e.g., by inadvertently affecting the tone or direction of the discussion).

VII. QUANTITATIVE SURVEY METHODS

Quantitative methods can be administered to large samples and can provide quantitative estimates for the outcomes of interest. Generally, a quantitative method appropriate for most TPPI studies is the self-report survey, although you should consider the strengths and weaknesses of various study methods to select the method most appropriate for your study aims. Surveys are a method of collecting data from participants by asking structured questions, typically with predetermined answer options, and can be implemented in different types of study designs. For instance, surveys can ask participants to provide numeric ratings of their intentions to try a product. These values allow for quantitative comparisons, including between tobacco products, between groups (e.g., users and nonusers), and between experimental conditions (e.g., to examine the effect of exposure to a modified risk claim on intentions).

A. Quantitative Survey Method Considerations

FDA recommends you follow certain practices for creating survey instruments to help reduce participant dropout and reduce measurement error by increasing the likelihood that participants

[9] *See, e.g.,* Krueger, R. A. (2014). *Focus groups: A practical guide for applied research*. Sage publications.
[10] *See e.g.*, Creswell, J. W., & Poth, C. N. (2016). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Sage Publications.

understand how to answer each question and reduce the likelihood that questions are written in a way that could bias responses. In particular, when developing survey instruments for a TPPI study, FDA recommends you:

- Present information to participants written at a reading level appropriate for those with less than a high school education;
- Avoid technical terms, and include definitions of terms that participants may not be familiar with or are likely to be misunderstood by participants;
- Include images of the product(s) when possible for questions referring to tobacco products other than combusted cigarettes;
- Keep the direction of the response scale consistent throughout the survey such that, for example, the most affirmative response options (e.g., *Definitely Yes* or *Strongly Agree*) are consistently at the beginning (or consistently at the end) of the response scale throughout the survey;
- Avoid including instructions or questions that contain information that educates participants or influences participants' ability to answer subsequent questions (e.g., reminding participants that a product contains nicotine in instructions and then asking a comprehension question about whether the product contains nicotine); and
- Consider ways to minimize order effects (i.e., differences in participants' responses due to the order in which study materials are presented to them). For example, investigators should consider the following:
    - Impact of previous questions or tasks on how participants may respond to subsequent questions;
    - Proximity of exposure to the stimulus and the primary outcome questions, as the effect of the stimulus may fade over time; and
    - Impact of any study tasks (instructions or measures) that precede exposure to the stimulus and how they may affect participants processing of the stimulus.
- Conduct cognitive interviews, where feasible and appropriate, before collecting the survey data. Cognitive interviews involve asking participants to explain what survey questions mean to them and the process they use to answer them.[11] This technique can detect potential problems with how participants understand, interpret, and answer each survey question. Potential problems might include items on the survey or response options that may be confusing or misinterpreted. FDA recommends refining the survey based on results of this cognitive interview testing and conducting additional rounds of testing on subsequent versions of the survey as necessary.[12]

Additionally, prior to launching a full-scale survey study, FDA recommends you:

- Conduct a pre-test by administering the survey to a small sample of participants, and carefully reviewing the data. Pre-tests can help reveal problems arising from the

[11] Willis, G. (2005). Cognitive Interviewing: A Tool for Improving Questionnaire Design. Thousand Oaks, CA: Sage.; Miller, K., Chepp, V., Willson, S., & Padilla, J. L. (Eds.). (2014). *Cognitive interviewing methodology*. John Wiley & Sons.

[12] Additional discussion of how to conduct cognitive interviews to refine survey questions can be found in Office of Management and Budget, *Statistical Policy Directive No. 2 Addendum: Standards and Guide Lines for Cognitive Interviews*, May 10, 2016 (81 FR 70587).

sequencing of questions, or issues with administration (e.g., errors in the programming of surveys administered electronically).

B. Choosing a Type of Quantitative Study Design

Various study designs can be appropriate for quantitative TPPI studies, including observational and experimental studies, which may be cross-sectional or longitudinal. In experimental studies, researchers examine the effect of manipulating an experimental factor on an outcome. Conversely, in observational studies, researchers make an effort to avoid affecting participant behavior. The goal is to observe and collect data on areas of interest without directly influencing a participant. In addition, longitudinal studies employ continuous or repeated measures to follow participants in a study over prolonged periods of time. Alternatively, cross-sectional studies analyze data from participants at a specific point in time. FDA recommends you take into account the strengths and weaknesses of various designs and select a study design most appropriate for addressing your specific study aims. For example, an experimental TPPI study allows for the assessment of whether changes in one factor (e.g., changes in a product or its label, labeling, and advertising) could cause changes in an outcome (e.g., perceptions of the product's risks).

If your aims include examining differences between products (e.g., a new product that has a different flavor than an existing product, or differences in use instructions on product labeling), or the effects of marketing a product as a modified risk (e.g., the addition of a modified risk claim in an advertisement), FDA recommends you conduct at least one study that uses an experimental design. For example, if you seek to support an MRTPA by evaluating the impact of a modified risk claim on consumers' perceptions of the proposed MRTP, we recommend selecting an experimental design. This type of experimental design could help demonstrate whether and how the change to the product (or presence of modified risk information) could cause changes in the study outcomes. [13]

In Section VII.C. of this guidance, we review particularly important practices to consider when conducting TPPI studies using experimental designs.

C. Experimental Design Considerations

FDA recommends using manipulation checks and carefully determining the study conditions when determining experimental designs.

First, FDA recommends that TPPI study procedure include a manipulation check, which determines whether the manipulated factor (or independent variable), such as a modified risk claim, was noticed by the participants as intended. Such an assessment is important for evaluating internal validity of the study. For instance, if the manipulation check reveals that participants did not notice the manipulated variable, then this could account for a failure to find effects on the study outcomes. There are a number of ways to implement a manipulation check.

[13] Note, this is conditioned on the assumption that the stimuli were presented in a way that enabled participants to notice and read key information (e.g., modified risk claim). This is further described below in Section VIII.

For instance, investigators may ask participants to recall (via free response) what they have just seen, or participants may be asked to select the target (e.g., modified risk information presented) from a list of options.

Prior to data collection and analysis, FDA recommends you consider how responses to the manipulation check will be used. For instance, you should determine whether these responses will be used to eliminate participants from the analytic sample, included as a covariate (control variable) in analyses, or used to conduct a sensitivity analysis (i.e., analysis to determine whether and how the exclusion of people who fail the manipulation check affects results). Because a manipulation check is used to evaluate a study's internal validity, an application referencing an experimental study without a manipulation check should include a scientific rationale for that decision.

Second, in determining the study conditions you will include, consider the comparisons you need to make to address your study aims. For instance, for a study assessing the effect of a modified risk claim in an advertisement, you generally would have at least one condition where participants are exposed to a stimulus (e.g., advertisement with the modified risk information being tested); and you would have at least one condition that is an appropriate control (e.g., the same stimulus without the modified risk information). Because the two conditions differ only in the presence (or absence) of modified risk information, comparisons between the two groups would enable inferences about the effect of the addition of the modified risk claim on the study outcomes.

VIII. MODE OF DATA COLLECTION

Each particular mode of data collection (e.g., online, phone, in-person) has its own strengths and weaknesses. FDA recommends that you take them into account and select the most appropriate mode of data collection for your study aims, design, method, and study population.
If you use online modes of data collection, FDA recommends you take steps to mitigate threats to validity that might arise from this mode, such as participants not paying attention or engaging in other activities while completing the study. These steps might include:

- Using procedures to ensure participants are complying with study instructions,
- Taking data quality control precautions (e.g., using well-defined criteria, established before the study is conducted, to identify and eliminate fraudulent cases),
- Carefully considering the display of study stimuli so that it will be legible and entirely visible on the devices through which data is collected.

FDA recommends that stimuli be presented in a way that ensures participants are able to see and read the stimuli. Correspondingly, FDA recommends that you consider the implications of a study's mode of administration (e.g., online vs. in-person) in determining the appropriate size and visibility of the stimuli. For example, studies conducted online should either ensure that participants using devices with small screens can appropriately view the stimuli or require participants to use a computer or other device with a screen of sufficient size to adequately view the stimuli. Relatedly, FDA recommends that your application include a detailed description of the procedures used to present stimuli to participants. This could include whether you required

participants to view the stimuli for a minimum length of time or whether the stimuli remained visible to participants throughout the study. FDA recommends that you provide a representation of the stimuli exactly as viewed by participants in the study. In addition, FDA recommends that you utilize stimuli in your study that are representative of the type of label, labeling, or advertising you may use in marketing the product.

If the product or its label, labeling, or advertising depicted in the stimuli is visually distinct from what you will submit in an application, FDA recommends you have a scientific rationale for why the study results are generalizable to what is submitted in an application. Also, FDA recommends that your stimuli reflect, to the extent possible, how consumers would view the product that is the subject of the application in the real world. For example, if your product packaging or advertising is required to bear warning statements, your study stimuli should include such warning statements.[14] FDA recommends including each required warning statement if the packaging or advertising is required to bear rotating warnings. In addition, FDA recommends that the stimuli include the same label, labeling, and advertising being proposed, as the context of the information can affect consumers' perceptions and understanding. Similarly, if you are seeking to market your proposed modified risk tobacco product with different combinations of modified risk information, FDA recommends that the study stimuli be designed to reflect such combinations. Finally, FDA recommends that studies testing the effect of a modified risk claim include stimuli for: (a) the experimental condition(s), which should include all of the modified risk information and any additional text to be used in advertising and labeling, and (b) the control condition(s), which should be identical to the experimental condition stimuli, with the exception of the removal of the modified risk information (or a replacement of the modified risk information with text, such as ad copy, that is not modified risk information).

IX. QUANTITATIVE MEASURES

Below are recommendations related to selecting, developing, and adapting quantitative study measures.

A. Selection and Development of Measures

It is important that TPPI studies use valid measures of study constructs, as the utility of the study's findings depends on the use of valid measures. The "validity" of a measure of a construct refers to the extent to which the variation between respondents' observed scores on the measure reflect the actual variation between respondents on the construct. For example, if a measure of absolute risk perception is valid, it means that individuals who perceive a product is low risk will have low scores on the measure, and individuals who perceive a product is high risk will have

[14] For cigarette packages and advertisements, the effective date of the final rule "Required Warnings for Cigarette Packages and Advertisements," is Oct. 16, 2021. The health warnings requirements for cigars and pipe tobacco is postponed until 60 days after final disposition of the plaintiffs' appeal of the court's order on the health warnings requirements. *See* Order, Cigar Ass'n of Am. v. U.S. Food & Drug Admin., No. 1:16-cv-01460 (D.D.C. July 5, 2018). FDA encourages applicants considering development of a study with warning statements for these products to discuss what would be appropriate for your product as part of the meeting recommended in section II of this document. See the Meetings Guidance in footnote 5 for more information on requesting meetings with FDA.

high scores on the measure. For measures of tobacco product perceptions and intentions, because the constructs are not directly observable, researchers typically assess a measure's validity using several methods, including by examining whether it is associated with other constructs in expected ways. For example, a valid measure of intentions to try a tobacco product could predict subsequent trial of that product. FDA recommends you select measures that have demonstrated some type of measurement validity (e.g., convergent validity, predictive validity) in peer-reviewed literature whenever possible and adapt them, as appropriate, for your study. For example, if you plan to measure intentions to quit smoking, the Motivation to Stop Smoking Scale would be an appropriate measure, as it has demonstrated predictive validity.[15] Alternatively, applicants could consider selecting measures that are widely used in the peer-reviewed literature, even if their validity has not been directly studied. In this case, peer-reviewed literature provides a basis for FDA to have more confidence in the validity of such measures. Additionally, if you choose to develop new measures, FDA recommends that you conduct research to assess the new measures' validity before use in the TPPI study. Whether you select measures from the literature or develop them, FDA recommends you have a scientific rationale for using each measure, as this can help provide support for the validity of your study findings.[16] FDA recommends the following additional guidelines for writing and adapting measures for TPPI studies.

B. General Recommendations for Writing or Adapting Measures

FDA recommends that measures of tobacco product perceptions, understanding, and intentions be written or adapted in a manner that specifically refers to the product (by name) that is the subject of the study. For example, a TPPI study concerning a cigarette product should ask respondents about their intentions to smoke the product that is the subject of the study (e.g., by sub-brand). Using measures that refer to the specific product by name maximizes FDA's ability to draw conclusions from your study findings about the tobacco product that is the subject of the application and informs FDA's review of the overall application.

In addition, it may be necessary to change certain aspects of a validated measure to better fit the product or study sample. Such aspects could include the measure's response scale, the specific domains assessed by the measure (e.g., the specific health outcomes assessed in a risk perception measure), or the reading level of the measure. When developing new measures, FDA also recommends you conduct cognitive interviews about your measures as described in Section VI.C of this guidance.

FDA has additional recommendations to consider when writing new measures or adapting existing measures for a TPPI study. These recommendations may help reduce the likelihood that characteristics of the measures will bias responses.

- Assure that each item is direct, specific, and unambiguous. Each item should address a single issue. Avoid compounded items that combine two or more questions into one

[15] Kotz, D., Brown, J., & West, R. (2013). Predictive validity of the Motivation To Stop Scale (MTSS): a single-item measure of motivation to stop smoking. *Drug and Alcohol Dependence*, *128*(1), 15-19.

[16] *See e.g.*, Dillman, D. A. (2011). Mail and Internet Surveys: The Tailored Design Method--2007 Update with New Internet, Visual, and Mixed-Mode Guide. John Wiley & Sons.

question. For example, the question "If you used this product, how likely are you to get lung cancer and heart disease?" should be broken into two questions, one that asks about lung cancer and another that asks about heart disease.

- Similarly, ensure that response scales assess only one dimension per item. For example, all points of an item response scale ranging from *Strongly Disagree* to *Strongly Agree* should assess disagreement and agreement only, and they should not assess a different dimension at the midpoint, such as "I don't think the advertisement affected me."
- Avoid leading questions, which are questions that suggest to the participant that the researcher desires a certain answer. For example, avoid questions such as "Now that you see how this product can improve your health, how likely are you to start using this product?"
- Avoid language in the question stem or instructions that could bias responses. For example, rather than stating "Please rate how strongly you agree with each statement," state "Please rate how strongly you agree or disagree with each statement"
- For questions that offer scale response options, include an appropriate number of response options. As part of determining this, consider that a greater range of response options, for example having more than three response options, allows for greater sensitivity to detect differences between conditions, if they exist. Cognitive interviews and pre-testing can help determine the appropriate number of response options for each item before finalizing the survey.
- When providing response scales, offer equal numbers of positive and negative options.
- When providing an 'undecided' or 'don't know' response option, distinguish these response options from 'neutral' by placing 'undecided' or 'don't know' options visually separate from the scale.

X. STUDY OUTCOMES

When writing your study protocol (i.e., prior to data collection), FDA recommends that you identify your TPPI study outcomes, based on your study hypotheses and research questions. While the relevant outcomes will vary according to the application type and the product that is the subject of the application, in general, the following are a list of outcomes that can be informative to questions FDA considers when reviewing applications.

A. Tobacco Product Perceptions

Consumer perceptions can relate to future behavior and can help inform FDA's evaluation of your application. There are numerous product perceptions that can be assessed. Determining which perceptions are most relevant and informative may depend to some extent on the applicable premarket authorization standard and type of tobacco product that is the subject of

your study.[17] In general, the following are perceptions that can be informative to the questions FDA addresses when reviewing applications; however, applicants may address additional perceptions:

- Perceptions about the absolute health risks of specific tobacco-related diseases, including addiction and other principal diseases associated with use of the product (e.g., heart disease)
- Perceptions about the health risks of using the product relative to:
  - Using other products in the same product category or using products in a different product category;
  - Using cessation aids or nicotine replacement therapy; and
  - Quitting all tobacco use.
- Perceptions of the health risks of dual use of the product and the comparison product, relative to both exclusive use of the proposed product and exclusive use of the comparison product.

When assessing product perceptions, FDA recommends the following practices:

- Prioritize asking questions in first-person (i.e., asking about participant's risk to themselves), because people sometimes engage in "unrealistic optimism" such that they acknowledge health risks to others while downplaying the extent to which those risks apply to themselves. Measuring perceived risks to oneself rather than to others could better predict future behavior.
- Specify the conditions of tobacco use, such as frequency and duration of use, so that participants are clear about the health behavior they are rating (e.g., smoking 10 cigarettes per day for the rest of their life). This can help reduce random error caused by differences in how participants interpret items.
- Ask about a variety of specific health risks and potential outcomes of tobacco product use to avoid missing key facets of product health risks (e.g., oral effects such as gum disease and mouth cancer, respiratory effects such as lung cancer, cardiovascular effects such as heart attack, addictive effects such as not being able to quit using the product).
- Use Likert-type response scales (i.e., scales using descriptive labels to represent a gradient of likelihood, severity, or agreement) rather than 100-point or other numeric response scales (e.g., the percentage or number of product users out of 100 who will experience a particular health effect), given that participants may have low numeracy and may find it difficult to express their perceptions on a numeric response scale.

[17] Additional information on topics of interest for the different premarket review pathways can be found in the respective guidances for each pathway (see footnotes 3-4) and, for additional information on MRTPs, see: Institute of Medicine (US). Committee on Scientific Standards for Studies on Modified Risk Tobacco Products. (2012). *Scientific standards for studies on modified risk tobacco products*. National Academies Press.

B. Behavioral Intentions

Behavioral intentions include, for example, intentions related to product purchase, trial, use, and discontinuing use. Intentions to use tobacco products are considered proximal predictors of behavior, and can help inform FDA's evaluation of your application, for example, by providing information relevant to evaluating how your product could affect the likelihood of use among different groups.

In general, FDA recommends you prioritize assessing the following behavioral intentions:

- Tobacco users' intention to try the product
- Tobacco users' intention to use the product regularly, and
- Non-tobacco users' intention to try the product, including intentions among never users (i.e., trial) and former users (i.e., relapse).

In addition to assessing the extent to which a participant intends to try the product that is the subject of the TPPI, behavioral intentions may also be assessed to determine how a potential user would be likely to use the product (i.e., expected patterns of use). For instance, the behavioral intentions may be assessed to determine whether a current tobacco user is likely to use the product exclusively, or to use as a complement to their current product use and therefore become a dual user. Alternatively, the behavioral intentions may be assessed to determine whether a current user is likely to entirely replace their current product with another product—for instance, entirely replacing combusted cigarettes with a non-combusted product. While a TPPI study may address how a participant intends to use the product, participants may have limited ability to forecast their future patterns of use behavior without having tried the product. Accordingly, patterns of use may be better assessed with data from behavioral studies, including actual use studies. Likewise, behavioral study data can address whether the product affects a tobacco user's likelihood of quitting tobacco completely (e.g., they start using a new product instead of quitting their current product).

C. Consumer Understanding

Consumer understanding can entail the understanding of information, such as instructions for use and principles of operation (e.g., instructions for recharging a battery), as well as understanding of modified risk information. For example, understanding instructions for use and principles of operation are important for products that require the user to perform certain tasks to use the product as intended (e.g., recharging a battery). Understanding is also important in the context of modified risk information. Consumer understanding could include assessing the extent to which participants:

- understand the health risks or exposure that are described as reduced when using the product as intended by the manufacturer;
- understand that other key health risks or exposures not described are not reduced when using the product as intended by the manufacturer; and
- understand the conditions of use for the product that are required to achieve reduced risk or reduced exposure (e.g., exclusive use).

To assess understanding of information presented to consumers, generally either new measures are developed, or existing measures are adapted to the specific information being assessed. For instance, information about modified risk or instructions for use is typically product-specific. Thus, to assess understanding of this specific information, applicants can either develop their own measures to assess this specific information or adapt measures from the literature to the specific information being assessed. We recommend applicants assess understanding with more than one type of measure and when designing these measures, consider the following, as appropriate:

- Including at least one measure with items that can be scored as "correct" or "incorrect";
- Including a measure that uses a hypothetical scenario, or 'vignette', to frame the question. As above, the vignette should specify the conditions of tobacco use, such as frequency and duration of use. For example, a measure could describe a person and his or her tobacco use behavior (e.g., dual use of smoking 10 cigarettes per day while also using the new product with a specified frequency) and then ask participants to rate the person's risk (e.g., risk of different tobacco-related diseases) after a certain time period (e.g., after 10 years); and
- Constructing multiple choice items (if used) carefully to avoid creating items that are too easy to answer, and thus less informative. For instance, avoid distractors that are obviously incorrect.

For example, in a study supporting an MRTPA, measures assessing understanding of a modified risk claim that exclusive use of the product lowers the risk of lung cancer compared to other products could include: the extent to which participants understand that lung cancer risk is reduced; the extent to which participants understand that key disease risks not included in the claim are *not* reduced; and the extent to which participants understand that they must use the product exclusively (i.e., they cannot engage in dual use) to achieve reduced risk of lung cancer.

As noted in Section IX.A above, FDA recommends that, when developing these measures, applicants should be able to support their validity—e.g., demonstrate that participants who correctly answer the questions are doing so because they understand the modified risk information, rather than because they can guess the correct answer without even viewing the modified risk information.

XI. QUANTITATIVE STUDY SAMPLE

FDA recommends that you determine each population that you want represented in your study based on your study aims and develop your sampling procedures with the goal of maximizing how well each sample represents those populations. Using samples that are representative of that population improves the generalizability of the results. For example, if a study aim is to determine likelihood of use by never smokers in the United States, the sampling procedures should be developed with the goal of maximizing the sample's representativeness of U.S. never smokers, so that study results can be generalized to U.S. never smokers. It is also important to clarify whether you are trying to generalize a prevalence estimate or an experimental effect. The relationship between two variables may be less likely to vary across samples than the absolute

prevalence of a position on a variable (e.g., prevalence of a belief or behavior). Below, we describe additional considerations regarding study samples.

A. Participant Sampling and Recruitment

FDA recommends that you consider how particular recruitment procedures could introduce bias that could affect how representative of the population your samples are, and therefore how well your results generalize to the population. For example, people of low socioeconomic status may be underrepresented in samples recruited online, which could limit the generalizability of the study's findings to the U.S. population. Therefore, in determining the study population, FDA recommends that you consider how you can minimize bias to maximize sample representativeness when:

- Locating potential research participants (e.g., random digit dialing, using an online panel);
- Inviting people to participate (what type of study recruitment materials such as emails or advertisements you will use to solicit participation in the study, how you will determine whom to invite from a pool of potential participants);
- Developing eligibility criteria;
- Determining which subpopulations (if any) you will oversample (see Section XI.B. for more information on oversampling); and
- Documenting the response and completion rates at the participant and item levels and accounting for missing data.

B. Populations of Interest: Users and Non-Users of Tobacco Products

To evaluate the potential impact of marketing a product, you may want to consider including in your study current users of tobacco products as well as people who do not currently use tobacco products (both former and never users). When identifying which user and nonuser groups to include in a study, you should consider: (1) your main rationale for how the study will help support that the product meets the statutory standards of the applicable pathway and (2) your study aims and hypotheses, product type, intended users, and unintended users of the product.[18] For example, consider a PMTA for a smokeless tobacco product that an applicant seeks to support with TPPI studies. These studies might include nationally representative samples of adults and might also oversample current smokeless tobacco product users (intended users) and young adult never tobacco users (unintended users), who might be more likely to experiment with a new tobacco product as compared to other demographic groups. Oversampling typically involves recruiting a disproportionately larger sample of a subgroup from the population, and results in increased statistical power to examine whether these groups differ from other subgroups or the main sample.

[18] For more information on the premarket review pathways and the MRTP pathway, see footnotes 3-4; For additional information on MRTPs and TPPI studies, see: Institute of Medicine (US). Committee on Scientific Standards for Studies on Modified Risk Tobacco Products. (2012). *Scientific standards for studies on modified risk tobacco products*. National Academies Press.

Once you have selected these user and nonuser groups, we recommend that you carefully consider what criteria you will use to define who belongs to each group. When determining criteria for user and non-user groups, FDA recommends you consult relevant scientific literature to determine what criteria are typically used to define group membership.

In all cases, FDA recommends that you clearly articulate how you are defining the various types of user and non-user groups (e.g., former users, never users) and have a scientific rationale for these definitions. Additionally, FDA recommends that you consider the potential impacts to vulnerable populations and whether you should oversample subpopulations that are especially likely to be affected by the marketing (including the label, labeling, and advertising) of the tobacco product being studied. Vulnerable populations may include, for example, groups that are susceptible to tobacco product risk and harm due to disproportionate rates of tobacco product initiation, use, or cessation, or burden of tobacco-related diseases. For example, young adults are typically more likely to initiate use of tobacco products relative to older adults, and thus may be particularly receptive to the marketing of a product.

C. Sample Size and Power

The sample size of a study is related to the study's statistical power. FDA recommends that you consider the following guidelines to determine your study sample size:

- If the study is quantitative with primary hypotheses, you should conduct statistical power analyses to determine the sample size needed to detect the hypothesized effect size(s). You should have a scientific rationale for your sample size that includes the following:
  - Statistical computations to determine sample sizes, specifying the number of primary hypotheses or research questions, the associated Type I error[19] probabilities, and the statistical power;
  - Study design and sampling plan; and
  - Statistical tests planned for analyses and the expected effect sizes for which the study was powered.
- If additional analyses are conducted on secondary hypotheses, you should consider that any lack of observed effects may be due to insufficient statistical power. You should also consider this when interpreting the results of quantitative exploratory studies (i.e., quantitative studies that are not designed to test specific hypotheses).
- FDA recommends that you develop a sampling plan, which would include your determination of sample size, to ensure, for instance, that you can detect modest effect size differences between vulnerable populations and the general population (see Section XI.B. of this guidance).
- For qualitative studies, FDA recommends that you determine the sample size using established guidelines for the design of the qualitative study being conducted. For example, one method for focus group and interview studies is to collect data until reaching *saturation* (the point at which continuing to collect data would no longer yield new information related to research questions). As with quantitative studies,

[19] A type I error (alpha error), generally, is the false identification of a change or effect that is not present.

FDA recommends you have a scientific rationale for how you determined the sample size.

XII. ANALYSIS PLAN AND REPORTING STUDY RESULTS

FDA recommends you use a style guide to help you format and report your study results, such as the style guide published by the American Psychological Association.[20] Adherence to an established style for reporting results facilitates FDA review by helping to ensure the results are presented in a complete and clear manner.

FDA recommends you develop an analysis plan before the data are collected. The analysis plan should follow from the hypotheses or research questions (see Section IV. of this guidance). Developing an analysis plan beforehand helps to prevent bias and promote transparency. Below are additional recommendations for reducing bias and promoting transparency are described below.

Your analysis plan should direct the analysis after data collection. If you deviate from the analysis plan because of unforeseen circumstances, FDA recommends you document the deviation, clearly explain why the deviation was necessary, and provide a rationale for any new approaches you may have employed as a result of the deviation.

A. Qualitative Analysis

For study designs that involve qualitative data analysis, FDA recommends that you consider the following when creating an analysis plan:

- How summarization of results, including qualitative coding, will be conducted;
- How agreement between different coders (i.e., interrater reliability) will be assessed and reported; and
- How themes will be derived during data analysis (e.g., using a coding scheme, using computer software).

B. Quantitative Analysis

If the study is quantitative, FDA recommends that you consider the following when you create the analysis plan:

- Power analyses for primary hypotheses to ensure the sample size is sufficient for the planned analytical approach (see Section XI.C. of this guidance);
- How raw data will be converted to an analyzable dataset, including creation of any new variables;
- How data will be assessed for meeting statistical assumptions of the chosen data analytic techniques;
- Specific types of analyses you will use or conduct to address primary and secondary hypotheses and research questions;

[20] *See*, http://www.apastyle.org/

- How you will assess effect size;
- Covariates (control variables) you plan to include, and a scientific rationale for including them (or a scientific rationale for not using covariates);
- How you will handle missing data;
- If you weight your data, weighting procedures and scientific rationale for how they were determined (or a scientific rationale for not using them); and
- Scientific rationale for how you dealt with Type I error with multiple comparisons.

Dichotomizing or categorizing continuous data involves collapsing data (e.g., categorizing data collected on a 5-point response scale into two categories for analysis). This can affect results and introduce bias to your findings because it involves an information loss and results can change based on how you define the categories. If you choose to dichotomize continuous data, you should have a scientific rationale for doing so, which should be included in the analysis plan.