Guidance for Industry

Label Comprehension Studies for Nonprescription Drug Products

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 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 document contact Captain Laura Shay, PhD, RN, USPHS, at 301-796-0994.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

April 2009
OTC

Guidance for Industry

Label Comprehension Studies for Nonprescription Drug Products

Additional copies are available from:

Office of Communications
Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 51, rm. 2201
Silver Spring, MD 20993-0002
(Tel) 301-796-3400
http://www.fda.gov/cder/guidance/index.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

April 2009
OTC

TABLE OF CONTENTS

Guidance for Industry[1]
Label Comprehension Studies for Nonprescription Drug Products

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

The Food and Drug Administration (FDA) often requires sponsors to conduct label comprehension studies that are designed to evaluate proposed nonprescription drug product labeling. This guidance is intended to provide recommendations to industry on conducting label comprehension studies. A label comprehension study is a tool that can be used for assessing the extent to which consumers understand the information conveyed by proposed nonprescription drug product labeling and then apply this information when making hypothetical drug product use decisions. Data derived from a label comprehension study can identify areas on the label that would benefit from clearer or simpler presentation of important consumer information.

It is important to note that label comprehension study data do not predict consumer behavior (e.g., how consumers actually use a drug product). Drug product use and other behaviors are often evaluated in an actual use study. The label used in an actual use study should be tested in a label comprehension study beforehand to ensure that consumers understand the information on the label.

This guidance covers general principles related to the conduct of label comprehension studies and should not be considered a substitute for an FDA review of specific protocols. This guidance incorporates advice obtained from the September 25, 2006, meeting of the Nonprescription Drug Advisory Committee that considered issues related to analysis and interpretation of consumer studies conducted to support marketing of nonprescription drugs.[2]

[1] This guidance has been prepared by the Division of Nonprescription Clinical Evaluation and the Office of Biostatistics in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration.

[2] The transcript from the September 25, 2006, Nonprescription Drug Advisory Committee meeting is available at http://www.fda.gov/ohrms/dockets/ac/06/transcripts/2006-4230t.pdf.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances 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 guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Under the Federal Food, Drug, and Cosmetic Act (the Act), the FDA has the authority to require sponsors to conduct label comprehension studies. Section 502 of the Act states that a drug product shall be deemed to be misbranded if its labeling is false or misleading in any particular way (21 U.S.C. 352(a)). In addition, section 502 states that a drug product is misbranded if its labeling fails to bear adequate directions for use (21 U.S.C. 352(f)). Furthermore, a drug product is misbranded if any word, statement, or other information required by or under authority of the Act to appear on the label or labeling is not "in such terms as to render it likely to be read and understood by the ordinary individual under customary conditions of purchase and use" (21 U.S.C. 352(c)).

Section 505(d) of the Act requires adequate tests by all methods reasonably applicable to show that a drug product is safe for use under the conditions prescribed, recommended, or suggested in proposed labeling (21 U.S.C. 355(d)). In addition, section 503(b)(1) of the Act requires an assessment of whether a drug product is safe for use without a prescription (21 U.S.C. 353(b)(1)). Moreover, FDA regulations further require that labeling "state the intended uses and results of the product; adequate directions for proper use; and warnings against unsafe use, side effects, and adverse reactions in such terms as to render them likely to be read and understood by the ordinary individual, including individuals of low comprehension…" (21 CFR 330.10(a)(4)(v)). Regulations on the format and content requirements for nonprescription drug product labeling are contained in 21 CFR 201.66.

The development of a nonprescription label is often an iterative process that depends upon testing and re-testing as the label evolves. Label comprehension studies can assess whether literate and low literate individuals can understand a drug product label. Some of the circumstances under which the FDA might require a label comprehension study include:

- Before the approval of a new drug product for the nonprescription market
- When one or more new indications, a new target population, or a new strength are proposed for a marketed nonprescription drug product
- When a substantive labeling change has been proposed (e.g., a change in the directions, a new warning)
- When drug products with new active ingredients that have a proprietary name associated with other active ingredients are proposed

- When a company generates multiple proprietary names for drug products containing identical quantities of identical active ingredients to assess consumer understanding of label information

- When adequate consumer labeling for the drug product warrants inclusion of a package insert in which case comprehension of the insert may need to be tested

Sponsors desiring FDA advice and consultation on a protocol for a label comprehension study should submit the protocol to an existing investigational new drug or new drug application in the Division of Nonprescription Clinical Evaluation.

III. STUDY DESIGN AND CONDUCT

When designing and conducting a label comprehension study it is important to:

- Clearly state the purpose of the study

- Identify the communication objectives (the important concepts that need to be understood by the consumer)

- Enroll a demographically diverse population with varying levels of literacy

- When necessary, enrich the study with subjects who have specific characteristics that are relative or absolute contraindications to use of the drug product

- Specify a study design that meets study objectives and calculate the appropriate sample size

- Construct a questionnaire that targets the communication objectives

- Use test labeling as close as possible to the final drug product label

- Minimize factors that may contribute to a biased study (e.g., sampling, recruitment strategies, leading questions, interviews that bias the responses in a particular direction)

- Compare different versions of the label to study the effect of variations in wording and information location on comprehension.

Label comprehension studies can be open-label, uncontrolled trials. A parallel group study design should be considered if the proposed nonprescription label is to be compared to existing labels, or different versions of the proposed label. Such study designs should be considered an important part of the process of developing an optimal nonprescription label. Sometimes more than one study may need to be conducted to develop a well-understood label.

A small pilot study or focus group testing should be conducted before the larger label comprehension study. This initial step can provide information on consumers' perceptions and knowledge about a drug product and the critical messages necessary for safe and effective use. This initial step can also help refine the label before it is tested in a larger study.

A. Study Objectives

The study protocol should clearly state the communication objectives of the study. A label comprehension study can have many communication objectives, but the most important objectives should be identified *a priori*. The characteristics of the active ingredient and the drug product class under consideration should determine what is important for consumers to understand, and therefore drive the communication objectives.

1. Primary Communication Objectives

In general, primary communication objectives should reflect information contained on the label that has the greatest clinical significance (e.g., indications, contraindications, warnings). A target level of comprehension for all communication objectives should be determined *a priori*. In general, the target level of comprehension for primary communication objectives should be 90 percent or greater; the greater the clinical significance (e.g., an absolute contraindication), the higher the target level should be.

Depending on the drug product, a study can have more than one primary communication objective. The following are examples of primary communication objectives:

- Consumer understanding of the indications
- Consumer understanding of dose and dosing interval
- Consumer understanding of contraindication(s), warning(s), and drug interaction(s)
- Consumer understanding of when to stop using the drug product

2. Secondary Communication Objectives

Secondary communication objectives also should be specified *a priori* with their target level of comprehension. In general, the target level of comprehension for secondary communication objectives should be 80 percent or greater. These secondary communication objectives often address areas less critical to the safe and appropriate use of the drug product, such as general health information (e.g., *when using this product continue a healthy diet and exercise*).

3. Self-Selection

Self-selection is the decision a consumer makes to use or not to use a drug product based on reading the information on the drug product label and applying knowledge of his or her personal medical history. Testing for appropriate self-selection can be conducted in a separate self-

selection study or it can be one of the objectives in a label comprehension study or actual use study.

B. Study Population

The study should include all subjects who could potentially use the drug product, regardless of their age, sex, underlying medical conditions, and use of concomitant medications. The study should test label comprehension in a general population whether or not individuals express interest in using the drug product. If particular populations are of interest, the study enrollment can be enriched according to one or more factors (e.g., age, subjects with specific disease for whom use of the drug product may be contraindicated). Because nonprescription drug products are available for purchase without a learned intermediary, and since no drug product is administered in the study, exclusion factors should be minimal and should be justified in the study protocol.

Label comprehension studies also should enroll a low literacy cohort. This low literate population should represent a range of low literacy below an 8th grade reading level. Education level is not a reliable substitute for literacy testing. At screening, the sponsor can assess literacy levels[3] of the study subjects by administering a validated instrument such as the Rapid Estimate of Adult Literacy in Medicine (REALM)[4] test (REALM-Teen for testing adolescents)[5] or the Test of Functional Health Literacy in Adults (TOFHLA).[6] Investigators should receive training to properly administer the literacy test. The number of low literacy subjects in the study sample should allow for a meaningful statistical analysis and inference to be made for this subgroup of subjects. If the label being tested requires the ability to understand and interpret numbers (e.g., weight- and/or age-based dosing directions), numeracy testing also should be considered using a validated instrument.[7]

C. Statistical Considerations and Data Analysis

1. Primary Endpoints and Success Criteria

The study protocol should provide a clear definition of the primary endpoints (primary variables), along with a rationale for the selection. The primary endpoints should be directly

[3] We recognize differences between health literacy and literacy measures. The REALM and TOFHLA were designed as rapid screening tools that were validated against the Wide Range Achievement Test for literacy; therefore, use of these instruments to screen literacy levels within the context of health is appropriate.

[4] Davis, TC et al., 1993, Rapid Estimate of Adult Literacy in Medicine: A Shortened Screening Instrument, Family Medicine, 25: 391-395.

[5] Davis, TC et al., 2006, Development and Validation of the Rapid Estimate of Adolescent Literacy in Medicine (REALM-Teen); A Tool to Screen Adolescents for Below-Grade Reading in Health Care Settings, Pediatrics, 118 (6): e1707-1714.

[6] Parker, RM et al., 1995, The Test of Functional Health Literacy in Adults: A New Instrument for Measuring Patients' Literacy Skills, Journal of General Internal Medicine, 10: 537-541.

[7] There are a number of numeracy screening instruments being used; however, this is a growing field. The FDA is interested in information on the numeracy tests that are considered validated screening instruments.

related to the primary communication objectives. The primary endpoints should be the endpoints capable of capturing the most relevant and convincing data on consumer comprehension of the critical label elements.

Based on the clearly defined primary endpoints, the study protocol should also specify what criteria determine success for the study. These success criteria should be related to the predefined target level of comprehension for the primary communication objectives. The success criteria should be defined using the confidence interval approach. This approach allows consideration of variability of the study data. For example, if the study has only one primary endpoint and is designed to ensure a predefined target level of comprehension, then the study can be claimed as a success only when the lower limit of the two-sided 95 percent (or one-sided 97.5 percent) confidence interval for the comprehension rate is above the target level.

We recommend sponsors use the two-sided 95 percent confidence interval to estimate the comprehension rate (or failure rate) and to define the success criteria. This approach sets the type I error for one-sided tests (2.5 percent) at half the conventional type I error (5 percent) used in two-sided tests, and generally is used for confirmatory clinical trials.

Typically, the study results need to demonstrate success for all the primary endpoints; however, if there are multiple independent primary endpoints, sponsors should address the issue of multiplicity to ensure that the overall error rate is appropriate. The confidence levels used in the success criteria should be adjusted accordingly.

For studies with secondary communication objectives, the protocol also should define the corresponding secondary endpoints and the success criteria depending on the purpose of the analyses of the secondary endpoints.

2. Sample Size Considerations

The number of subjects in a label comprehension study should be large enough to provide a reliable answer to the primary communication objectives. Sizing of such a study should be based on the success criteria. This generally involves the predefined target level P_0 for the comprehension rate, the assumed comprehension rate P_1 for the study population, the type I error rate α, and the type II error rate β.

Typically, the type I error rate α is set at 2.5 percent. The type II error rate β can be range 10 percent to 20 percent. The target comprehension rates can vary depending upon the medical significance of communication objectives. For example, for primary communication objectives, the goal might be to aim for a target comprehension rate, such as a P_0 of 0.90 to 0.95.

If multiple primary communication objectives are evaluated independently in the study, then the sample size should be adjusted for the multiple confidence interval calculations for each of the primary communication objectives. The number of subjects in a label comprehension study should be large enough to evaluate the primary communication objectives for important subgroups, such as the low literate population.

3. Data Analysis

The principal features of the planned analysis should be clearly defined in the protocol. The statistical methods for characterization of study subjects, analysis of the primary and secondary endpoints, and safety data should be specified in the protocol. Methods for constructing a two-sided confidence interval to estimate and define the success criteria for the comprehension rate (or failure rate) should be described. Methods for handling missing data and multiplicity should be specified. In some circumstances, a separate and comprehensive statistical analysis plan should be provided to address all the details of the data analysis.

D. Questionnaire Design

The questionnaire design should: 1) clearly reflect the communication objectives of the study; and 2) optimize the validity and interpretability of the information collected. Wording, question structure, and question sequences significantly affect the validity and interpretability of the data collected. A detailed discussion of questionnaire development is beyond the scope of this guidance. We recommend that sponsors consult experts in questionnaire design. The following points merit particular consideration:

- Questions should be designed to assess the specific communication objective.

- Simple vocabulary and pretest questions should be used to ensure questions illicit the intended information.

- Questions should be direct, specific, and unambiguous. Each question should address a single item or issue.

- Questions should test whether subjects can apply the information on the label. For example: Jennifer's child is 8 years old and weighs 52 pounds. How many teaspoons of Drug X should Jennifer give him?

- Different types of questions should be used, such as open-ended or closed-ended; a combination of these types of questions is encouraged. Scenario questions that are based on hypothetical situations also can be used. For example: Sally is pregnant and would like to take Drug X. Is it okay or not okay for Sally to take Drug X?

- Closed-ended questions should be validated with an open-ended probing question, otherwise subjects have a 50 percent chance of being correct by chance alone. For example: John has diabetes and would like to take Drug X. Is it okay or not okay? Why did you say that?

- If subjects answer incorrectly, verbatim responses should be collected using open-ended probing questions to assess why they answered the question the way that they did. It is important to collect this information to determine what changes to the label are needed to improve comprehension.

- Biasing questions, such as leading questions, should be avoided. An example of a leading question is: Joe stopped taking Drug X and went to see his doctor because he developed a rash. Is this okay or not okay?

- Questions that may cause framing or mindset bias should be avoided. An example of this type of bias is providing the response category of *ask a doctor* for all multiple-choice questions. When they do not know an answer, study subjects are more likely to choose *ask a doctor* rather than *I don't know*.

- Questions should be ordered so that information contained in a question does not bias a subject's ability to answer subsequent questions.

- Response choices in multiple-choice questions should be mutually exclusive and independent and contain only one correct answer.

- When listing response categories for multiple-choice questions, the category *I don't know* should be included as one of the response categories to give subjects permission to admit that they do not know so they avoid guessing.

- If a label comprehension study includes testing the subject's ability to appropriately self-select, questions that are used to validate the self-selection decision should be asked at the end of the study. Prompting subjects to think about their medical history before they make a self-selection decision or are tested on label comprehension can bias the study.

- Questions intended to measure the behavioral intent of the subject should not be used. Testing behavior is outside the scope of a label comprehension study. An actual use study should be conducted if information about how subjects would behave under nonprescription conditions is needed.

- Pretesting the questionnaire with a sample of respondents similar to the target population to ascertain that the questionnaire is eliciting the intended information should be standard practice. Pretesting can provide an extremely useful validation procedure.

The following two general approaches to administering the questionnaire can be considered: 1) self-administration; and/or 2) asking the questions using a trained interviewer. Using a trained interviewer may lessen the chance that low literate subjects will incorrectly respond because they cannot comprehend the written question when, in fact, they understand the label. Using an interviewer, however, may lead to interviewer bias particularly if the interviewer leads the subject to elicit a response. Interviewers involved in the study should be adequately trained, and have standard protocols and/or scripts to adhere to, especially regarding questions that subjects might ask.

E. Label Versions and Format and Content Requirements

Sponsors should consider testing different versions of the label in different studies or testing and comparing comprehension of several variations of a label within the same study.

The standardized nonprescription Drug Facts Label format and content requirements should be used.[8] If a sponsor chooses to deviate from the Drug Facts Label (see 21 CFR 201.66(e) and (f)), a rationale should be provided and testing should be conducted comparing the deviation with the Drug Facts Label.

F. Study Conduct and Location

In an effort to reflect *customary conditions of purchase*, a label comprehension study generally should not be carried out in a clinical or simulated clinical setting. The study site can be in a mall, or in places frequented by consumers. It also can be designed to simulate an actual purchase site. The study setting should be comfortable and well lit for reading. Subjects should have adequate time to read the label and be able to refer to it throughout the testing period.

Subjects should receive sufficient instruction on the format and conduct of the study and the expected length of time it will take to participate. Well-trained study site investigators should carry out procedures according to the protocol. Investigators should adhere to scripted responses to subject queries.

G. Data Collection, Recording, and Auditing

Verbatim responses to all questions should be recorded. The procedure for coding, categorizing, and analyzing verbatim responses to open-ended questions should be specified *a priori* in the protocol. All correct and incorrect answers to closed-ended questions also should be prespecified.

Methods for verification of complete and accurate recording of study data (i.e., subjects' responses, data entry, missing data, and data coding) should be described in the protocol.

IV. FINAL STUDY REPORT

The final study report should summarize the study design, conduct, and interpretation of the study results. The demographic characteristics of the study subjects, including literacy level, should be presented in the study report.

Optimally, the study subjects should represent the target population. Therefore, the results of the sampling effort should be assessed to determine whether this goal has been attained. The study report should describe the nature of the recruitment effort and the response rate (i.e., the proportion of screened subjects who were actually enrolled in the study). If possible, potential subjects who were excluded or chose not to enroll in the study should be characterized by demographic factors and the reasons for nonparticipation. Enrolled subjects should be characterized as to relevant demographic factors and whether or not they completed the entire study. Reasons why subjects failed to complete the study should be provided in the study report.

[8] Format and content requirements for nonprescription drug product labeling can be found under 21 CFR 201.66.

The presentation of the study results should include both the overall comprehension rates and comprehension rates in appropriate subsets (e.g., literacy level, sex, age, race, and presence of high risk factors).

V. INTERPRETATION OF STUDY FINDINGS

The acceptable comprehension level of a communication objective should be based on meeting the success criteria established *a priori*. The interpretation of these quantitative data also should be supported by the verbatim responses collected for each of the communication objectives. There may be times when the quantitative information reflects correct comprehension but the verbatim responses do not and visa versa. Thus, a clear analysis of both quantitative and qualitative data types should be provided to support and interpret the study findings.