Guidance for Industry

Label Comprehension Studies for Nonprescription Drug Products

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

August 2010
OTC

GRZAN

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; Fax: 301-847-8714; E-mail: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

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

August 2010
OTC

GRZAN

TABLE OF CONTENTS

I. INTRODUCTION 1

II. BACKGROUND 2

III. STUDY DESIGN AND CONDUCT 3

A. Study Objectives 3

1. Primary Communication Objective 4
2. Secondary Communication Objectives 4
3. Self-Selection Objective 4

B. Study Population 4

C. Statistical Considerations and Data Analysis 6

1. Primary Endpoints and Success Criteria 6
2. Sample Size Considerations 6
3. Data Analysis 7

D. Questionnaire Design 7

E. Label Versions and Format and Content Requirements 9

F. Study Conduct and Location 9

G. Data Collection, Recording, and Auditing 9

IV. FINAL STUDY REPORT 9

V. INTERPRETATION OF STUDY FINDINGS 10

GRZAN

Contains Nonbinding Recommendations

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

This guidance represents 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) sometimes 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 assesses the extent to which consumers understand the information on nonprescription drug product labeling and then apply this information when making drug product use decisions in a hypothetical situation. 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. We recommend that the label used in an actual use study 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] and comments received regarding the draft guidance published on May 1, 2009.

[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 503(b)(1) of the Act requires an assessment of whether a drug product is safe for use without the professional supervision of a practitioner licensed by law to administer such drug product (21 U.S.C. 353(b)(1)). 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 under customary conditions of purchase and use" (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 an iterative process that may depend upon testing and re-testing as the label evolves. Label comprehension studies should 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) for a marketed nonprescription drug product
- When drug products with new active ingredients that have a proprietary name associated with other active ingredients are proposed

- When adequate consumer labeling for the drug product requires the inclusion of a package insert; comprehension testing of the insert may be needed

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:

- 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
- 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)

Label comprehension studies are open-label, uncontrolled trials. Qualitative research and pilot testing should be conducted comparing different versions of the label prototypes before the larger label comprehension study is conducted. This enables refinement of the study questionnaire and the label before the label is tested in a larger study.

A. Study Objectives

The goal of a label comprehension study should be to test consumer comprehension of the major communication messages that detail the safe and effective use of a nonprescription drug product. These major messages should be the communication objectives for the study. Other messages on the label do not necessarily need to be tested.

The study protocol should state the communication objectives. A label comprehension study can have many communication objectives. All the communication 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 Objective

The primary communication objective should be the major communication message with the greatest clinical consequence to the consumer. Depending on the drug product, a study can have more than one primary communication objective. A target level of comprehension for each primary communication objective should be determined a priori. The target level of comprehension should reflect the clinical significance of the primary communication objective.

Depending upon the drug product, the primary communication objective can address:

- Consumer understanding of the indication
- Consumer understanding of dose and dosing interval
- Consumer understanding of specific contraindications, warnings, and drug interactions
- Consumer understanding of when to stop using the drug product

2. Secondary Communication Objectives

Secondary communication objectives also should be specified a priori. Not all label comprehension studies will have secondary communication objectives. 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*). Prespecified target comprehension levels are not needed for secondary communication objectives.

3. Self-Selection Objective

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. It is optimal to test for appropriate self-selection in a separate self-selection study or in the setting of an actual use study. However, in some circumstances it may be useful to test for appropriate self-selection as part of a label comprehension study. If this is the case, sponsors should enrich the study with subjects who would be interested in using the drug product, and subjects with specific relative or absolute contraindications to use of the proposed drug product, to ensure that the population at greatest risk understands the label.

B. Study Population

The study should include all subjects who can potentially use the drug product, regardless of 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. 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 (e.g., inability to read and understand English) and should be justified in the study protocol.

Label comprehension studies should also include an adequate number of subjects who have low literacy skills to examine the comprehension of the label in this subgroup. The proportion of low literacy subjects in the study sample should be representative of the proportion of adults in the United States with basic literacy skills based on available national data. The average reading level in the United States is estimated to be 8th grade.[3] Standard practice is to write medical information at a 4th to 5th grade reading level.[4,5] Therefore, attempts should be made to write the nonprescription label at a 4th to 5th grade reading level and no higher than an 8th grade reading level.

To adequately test the label, the low literate subjects should consist of an equal distribution of consumers who have 4th to 8th grade reading skills or marginal functional health literacy skills. Education level is not a reliable substitute for literacy testing. At screening, the sponsor should assess literacy levels of the study subjects by administering a validated instrument such as the Rapid Estimate of Adult Literacy in Medicine (REALM) test,[6] REALM-Teen for testing adolescents,[7] or the Test of Functional Health Literacy in Adults (TOFHLA or S-TOFHLA).[8,9,10] Investigators should receive training to properly administer the literacy test. If the label being tested requires the ability to understand and interpret numbers (e.g., weight- and/or age-based dosing directions), it may be appropriate for sponsors to screen for functional health literacy using the TOFHLA or S-TOFHLA.[11]

---

[3] Cotunga, N, CE Vickery, and KM Carpenter-Haefele, 2005, Evaluation of Literacy Level of Patient Education Pages in Health-Related Journals, Journal of Community Health, 30(3), 213-219.

[4] Andrus, MR and MT Roth, 2002, Health Literacy: A Review, Pharmacotherapy, 22(3), 282-302.

[5] Paasche-Orlow, MK et al., 2003, Readability Standards for Informed-Consent Forms as Compared With Actual Readability, New England Journal of Medicine, 348:721-6.

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

[7] REALM-Teen: 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.

[8] TOFHLA: 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.

[9] Baker, DW et al., 1999, Development of a Brief Test to Measure Functional Health Literacy, Patient Education and Counseling, 38:33-42.

[10] 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.

[11] TOFHLA: 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.

C. Statistical Considerations and Data Analysis

1. Primary Endpoints and Success Criteria

The study protocol should specify primary endpoints (primary variables), along with rationales for their selection. Primary endpoints should directly relate to the primary communication objectives. 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 specified primary endpoints, the study protocol should also specify success criteria (i.e., what determine success for the study) that 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, because it considers uncertainty in the sample 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 predefined 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 rate for one-sided tests (2.5 percent) at half the conventional type I error rate of 5 percent used in two-sided tests.

Typically, label comprehension studies have multiple primary endpoints and should demonstrate success for all the primary endpoints. In this case, the primary endpoints are referred to as co-primary endpoints. If the endpoints are not co-primary endpoints, the study protocol should address multiplicity issues to ensure that the overall type I error rate is appropriate and the confidence intervals are adjusted accordingly.

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 for the comprehension rate, the assumed comprehension rate for the study population, the type I error rate, and the type II error rate (or the study power).

The type I error rate should be set at 2.5 percent. The type II error rate can be in the range of 10 percent to 20 percent. The target comprehension rates can vary depending upon the medical significance of the communication objectives.

If the primary endpoints are not co-primary, then the sample size should be adjusted for the multiple confidence interval calculations for each of the primary endpoints. The number of subjects in a label comprehension study should be large enough to evaluate the primary endpoints for important subgroups, such as the low literate subgroup.

3. Data Analysis

The principal features of the planned analysis should be defined in the protocol. The statistical methods for characterization of study subjects, and the analysis of the primary and secondary endpoints 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 of the primary endpoints (or failure rate) should be described. Methods for handling missing data should be specified. A comprehensive statistical analysis plan should address all the details of the data analysis.

D. Questionnaire Design

The questionnaire design should reflect the communication objectives of the study and optimize the validity and interpretability of the information collected. Wording, question structure, and question sequences may 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 objectives.

- Simple vocabulary and pretested questions should be used. Pretesting the questionnaire with a sample of respondents to ascertain that the questionnaire is eliciting the intended information should be routine practice.

- 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 son is 8 years old and weighs 52 pounds. According to the label, how many teaspoons of Drug X should Jennifer give him?

- Different types of questions, such as open-ended, closed-ended, and multiple choice questions can be used.

- Scenario questions that are based on hypothetical situations can be used. Scenario-based questions are often used to assess the ability to make correct decisions based on information on the label. For example: Sally is pregnant and would like to take Drug X. According to the label, is it okay or not okay for Sally to take Drug X? Answers to the closed-ended question (e.g., *okay* or *not okay*) should be validated with answers to 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?

- 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? An example of a nonleading question is: Joe developed a rash after he started taking Drug X. What should Joe do?

- Questions that may cause framing or mindset bias should be avoided. An example of this type of bias is providing the response category of *talk to a doctor* for multiple-choice questions. If the correct answer to the question is *talk to a doctor*, a multiple-choice question with a response category of *talk to a doctor* may be leading. When a subject does not know an answer, he or she may choose *talk to a doctor* rather than *I don't know* because it seems like the right thing to do. This type of bias is often referred to as social desirability. In this situation, a multiple-choice question is not recommended.

- Questions should not contain information that educates and influences 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. A limited use of multiple-choice questions is recommended. Questions that require subjects to generate meaning or content on their own (rather than simply selecting an existing answer option) provide more reliable data to assess a subject's ability to comprehend written material.

- 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. The reasons for incorrect answers can also be assessed.

- 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.

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

- Other types of questions that have not been described that do not introduce bias may be appropriate.

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 comprehend the label. However,

using an interviewer 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. Sponsors should consider inherent bias that can occur with any data collection method. Therefore, they should provide their rationale as to why they chose a particular method and should address any potential bias.

E. Label Versions and Format and Content Requirements

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

F. Study Conduct and Location

Advertisements for the study should not contain information about the proposed drug product. Interested subjects should be provided only with information on how and where to participate in the study. The study site can be in a mall or in other places frequented by consumers. The study setting should be comfortable and well lit for reading. Subjects should be told before the study starts that they can refer to the label during questioning. Questions can begin with the statement *according to the label*; however, subjects should not be overly prompted to look at the label during questioning.

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. Correct and incorrect answers to closed-ended questions also should be prespecified. Post-hoc coding for open-ended questions should be documented.

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

IV. FINAL STUDY REPORT

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

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). Potential subjects who were excluded should be characterized by the reasons for exclusion. 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.

The presentation of the study results should include both the overall comprehension rates and comprehension rates in appropriate subgroups (e.g., low literacy, normal literacy).

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 where applicable. There may be times when the quantitative information reflects correct comprehension but the verbatim responses do not and vice versa. Thus, an analysis of both quantitative and qualitative data types should be provided to support and interpret the study findings.