Guidance for Industry

Clinical Studies Section of Labeling for Prescription Drugs and Biologics-- Content and Format

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.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be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance.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 Dockets Management Branch (HFA-305), Food and Drug Administration, 12420 Parklawn Dr., rm. 1-23, Rockville, MD 20857. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding the content of this draft document, contact (CDER) Janet Jones, 301-594-6758 or (CBER) Toni Stifano, 301-827-6190.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
May 2001
Labeling

Guidance for Industry

Clinical Studies Section of Labeling for Prescription Drugs and Biologics-- Content and Format

Copies of this Guidance are available from:

Office of Training and Communications
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Rockville, MD 20857
(Phone 301-827-4573)
Internet: http://www.fda.gov/cder/guidance/index.htm.

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
Internet: http://www.fda.gov/cber/guidelines.htm.
Fax: 1-888-CBERFAX or 301-827-3844
Mail: the Voice Information System at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
May 2001
Labeling

TABLE OF CONTENTS

Guidance for Industry[1]

Clinical Studies Section of Labeling for Prescription Drugs and Biologics--Content and Format

I. INTRODUCTION

This guidance is intended to help applicants decide (1) what studies should be included in the CLINICAL STUDIES section of medical product labeling, (2) how to describe individual studies, and (3) how to present study data, including presentation of data in graphs and tables. This guidance is intended to make the CLINICAL STUDIES section of labeling more useful to prescribers and to promote consistency in the content and format of the section across drug product classes and within drug classes and indications. This guidance also calls attention to the advertising and promotional implications of data and statements contained in the CLINICAL STUDIES section.

The overriding objective in labeling is to provide the information that is most useful to prescribers in treating their patients. In some cases, making the information in the CLINICAL STUDIES section of labeling more useful to prescribers could warrant significant departures from past labeling practices.

II. IDENTIFYING STUDIES FOR INCLUSION IN THE CLINICAL STUDIES SECTION

The CLINICAL STUDIES section of product labeling should provide a concise, accurate summary of the evidence supporting effectiveness — generally, the adequate and well-controlled

[1]This guidance has been prepared by the Medical Policy Coordinating Committees in the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

studies[2] that address effectiveness of the drug or biologic[3] for its approved indication. This section of the labeling is not intended to describe all available efficacy[4] data. Redundant information should be omitted or described briefly without detail. If there are multiple studies that address the same effectiveness issue, the subset selected for presentation should ordinarily reflect the overall conclusions derived from the application database as a whole (e.g., should not suggest a larger treatment effect than the database as a whole). However, study results that are inconsistent with the overall conclusions (e.g., absence of a treatment effect) should be included when they provide important information about drug effectiveness that is not otherwise available (e.g., information about a population subset, dose response, or the limitations of effectiveness).

A. Studies That Should Usually Be Included in the Clinical Studies Section

The following are the types of studies that should usually be included in the CLINICAL STUDIES section:

- Clinical studies that provide primary support for effectiveness.
- Clinical studies that provide important information about the limitations of effectiveness.
- Other clinical studies that contribute important efficacy data not provided by those trials that provide primary support for effectiveness (e.g., information about a population subset, dose, dose-response, effect size).

B. Studies That Should Usually *Not* Be Included in the Clinical Studies Section

The following are the types of studies that should usually not be included in the CLINICAL STUDIES section, unless they also meet one of the factors in II.A (above):

- Clinical studies with results that imply effectiveness for an unapproved indication.
- Active control clinical studies that imply comparative efficacy or safety claims not supported by substantial evidence.
- Studies that are not adequate and well-controlled. For any exceptions, the limitations of the study and the compelling reasons for inclusion should be stated.

---

[2] See 21 CFR 201.57(m).

[3] This guidance applies to drug and biological products. For the purposes of this guidance, the term *drug* will be used to include both drug and biological products.

[4] As used in this guidance, the term *efficacy* refers to the findings in an adequate and well controlled clinical trial, and the term *effectiveness* refers to the regulatory determination that is made on the basis of clinical efficacy and other data.

III. DESCRIBING STUDIES IN THE CLINICAL STUDIES SECTION

A. General Principles

1. Focus on Efficacy Data

The primary objective of the CLINICAL STUDIES section is to summarize (1) the evidence supporting effectiveness in the patients who were studied, (2) the critical design aspects of the studies, including the populations studied and endpoints measured, and (3) the important limitations of the available evidence. Ordinarily, safety data should be described in the ADVERSE REACTIONS section. However, in unusual cases it may be more clear or concise to present critical safety data in the CLINICAL STUDIES section, along with the study description and the efficacy data (e.g., when there is a study designed to evaluate a specific safety endpoint). If safety data are presented in the CLINICAL STUDIES section, they should be cross-referenced in the ADVERSE REACTIONS section and other sections, as appropriate.

2. Amount of Detail

In general, the amount of detail needed to provide a useful description of a study will depend on the indication, the trial design, the familiarity of the drug or drug class, and the extent to which the information adds to an understanding of the clinical effects of the drug and how the drug should be used.

Ordinarily, more detail is needed in the following situations:

- The study responses measured are of critical health importance. Typically, such responses would be direct measures of a meaningful clinical outcome (e.g., mortality, stroke or acute myocardial infarction rates, fracture rates) but could also include effects on important surrogate endpoints (e.g., cholesterol or hemoglobin A1C).

- The study results represent a significant advance in the treatment of a disease or condition or provide important information about a drug's activities relative to the drug's therapeutic class.

- The study enrolled a very specific population and the results may not be applicable to other populations.

- The study results are not typical of what would be expected for that drug class and indication. For example, the study results demonstrate a particularly marginal response or a response for which the clinical meaning or implications are unclear.

- The study results demonstrate that a new agent offers a clear advantage over existing therapy (see section III.A.4 below for discussion of comparative claims).

- The study uses an unfamiliar endpoint (e.g., a novel surrogate endpoint), or there are important limitations and uncertainties associated with an endpoint.

Ordinarily, less detail is needed in the following situations:

- The new drug appears to have effects that are typical of its class.

- The clinical endpoints measured in the study are not readily measurable or applicable in clinical practice (e.g., exercise testing in a study of heart failure can demonstrate effectiveness but does not translate to a measurable clinical outcome).

In these cases, it could be useful to describe the study in general terms (e.g., population, duration, measurements, and qualitative outcome) without providing detailed results.

3. Endpoints

The CLINICAL STUDIES section should present those endpoints that are essential to establishing the effectiveness of the drug (or that show the limitations of effectiveness) and those that provide additional useful and valid information about the activities of the drug. Endpoints presented should be endpoints the Agency has accepted as evidence of effectiveness, or closely related endpoints that may be more easily understood by clinicians. When it would be informative, the CLINICAL STUDIES section can also discuss other endpoints that were shown to be affected by the drug and endpoints that would have been expected to be influenced by the drug, but were not.

- **Composite Endpoints:** In general, effects on all components of a composite endpoint should be presented. Presentation of all components reveals which components are driving the result and which components may be unaffected, or even adversely affected, by treatment with the drug. When there is a range of effects on the components of a composite endpoint, selectively presenting only a single component of the composite endpoint, or presenting only the change in the composite endpoint, can be misleading.

- **Primary and Secondary Endpoints:** The terms *primary endpoint* and *secondary endpoint* should only be used when they would be helpful to understanding a drug's effect.

- **Closely Related Endpoints:** If two or more endpoints are closely related and convey essentially the same information, only one should be presented.

4. Comparative Data

Comparative data should generally not be included in labeling unless the data are from adequate and well-controlled studies that can support a comparative claim. If, however, the results from an active comparator arm and identity of the active comparator contribute information that is essential to a clinician's understanding of the drug's effects, the results and identity should be included in labeling. In such cases, the labeling should make clear that no comparative claim has been established and should disclose any limitations of the comparative data (e.g., if the comparator was administered in a suboptimal or unapproved regimen).

For example, when describing a clinical trial with three treatment arms (study drug, active control, and placebo) in which the comparison of study drug to placebo yields important efficacy information, the name of the active control and the results from that arm should be omitted if those data are not adequate to support a comparative claim. In contrast, when an active control, non-inferiority trial is critical to establishing effectiveness of a new drug, the name of the control and the results from the control arm should be included even though the data do not support a comparative claim. The labeling should indicate that the data do not support a comparative claim and should disclose any other limitations of the data.

B. Describing the Study Design

Usually, the description of the study design should include the following:

1. Major Design Characteristics

The major design characteristics should be identified, including level of blinding (e.g., double-blinded, partially blinded, open-label), type of controls (e.g., placebo, active, historical), duration of the study, method of allocation to treatment groups (e.g., randomization), and use of a run-in period to identify potential responders or eliminate placebo responders from subsequent phases of the study. Often these factors can be summarized in a phrase such as "randomized, double-blind, placebo-controlled study."

2. Treatment Arms

The dose, regimens, and any titration procedure should be identified for each trial arm.

3. Concomitant Therapy

Information about concomitant therapies should be included to the extent it is important to understanding the use of the study drug or its effects.

4. Study Population

The description of the study population should identify those characteristics of the population that are important to understanding how to interpret and apply the study results. The description should identify important inclusion and exclusion criteria, demographic characteristics, baseline values of any clinically relevant variables that would be important to understanding the treatment effect, and other characteristics of the population that have implications for the extent to which results can be generalized. For example, the description should discuss enrollment factors that exclude patients prone to adverse effects, the age range of the study population, a baseline value that results in a study population that is more or less sick than usual, or a study population enriched by a study design that eliminates nonresponders.

5. Critical Endpoints

Endpoints critical to establishing efficacy should be identified, and those that are not commonly understood should be defined.

C. Summarizing Study Findings

When a detailed summary of study findings is important to understanding the clinical effects of the drug (see section III.A.2 for a discussion of when more detail is important), the following elements should be addressed:

1. Disposition of Patients

Ordinarily the discussion of disposition of patients should include the following:

- The number of patients enrolled.
- The number of subjects completing the study.
- The number of patients discontinuing the study and reasons for discontinuation.
- For a study with a run-in period or other distinct phases, the number of patients entering each phase and the number of patients not progressing to the next phase (can be very important for understanding the study results).

2. Treatment Effect[5]

The summary of findings should describe the clinical outcome of the treatment relative to comparator (e.g., placebo or active).

- **Absolute vs. Relative Difference:** When presenting differences between study group and comparator, it is important to present the absolute difference between treatment groups for the endpoint measured, not only the relative difference. For example, if mortality is 6% in one study arm and 8% in the other, the absolute difference (2%) should be presented.

- **Group Results and Individual Patient Data:** Typically, the treatment effect is presented as a mean or median result accompanied by a measure of uncertainty or distribution of results for the treated groups. However, providing individual patient data for all treatment groups can be a useful alternative for describing the clinical effect of a drug. This can be done by including a graphical presentation of the distribution or cumulative distribution of responses among individual patients (see appendix for examples of graphical methods for presenting individual patient data).

- **Combined Data:** In certain situations, analyses of data combined from multiple efficacy studies can be useful for estimating the treatment effect. These analyses should be included only when they are scientifically appropriate and useful to better characterize the treatment effect. Meta-analytic graphs (see appendix) can be useful for displaying confidence intervals from several studies.

- **Uncertainty of Treatment Effect:** A confidence interval is typically more informative than a p-value and is the preferred method for describing uncertainty of the treatment effect. Although both a confidence interval and a p-value provide information about the uncertainty of the treatment effect, the confidence interval also provides information about the likely size of the treatment effect. A p-value can be included with a confidence interval, but should not be used alone, as it is potentially misleading.

3. Describing Results Within Treatment Groups

Because the comparison between treatment groups is critical to an understanding of the treatment effect, results for both the study drug and comparator should be presented. There is almost never a reason to show only results from the study drug group. This is especially important for studies with large effects in the placebo group, where presentation of results uncorrected for the placebo group

[5] *Treatment effect* means the effect that can be attributed to the drug. It is typically derived from a comparison of two prospectively identified treatment arms. Examples of such comparisons include differences in proportions, differences in means, or hazard ratios.

response can be highly misleading. When results from active control arms are discussed, caution should be taken to avoid the implication of an unsubstantiated comparative claim (see section III.A.4). The presentation of results within a treatment group should include, where appropriate, information about the variability of individual patient responses within the treatment group. This can be accomplished with, for example, standard deviations or box plots (see appendix for examples of graphical methods for presenting results within treatment groups).

4. Demographic Subgroups

The CLINICAL STUDIES section should include a summary statement about the results of required explorations of treatment effects in age, gender, and racial subgroups (21 CFR 314.50). Compelling results from analyses of other subgroups of established interest should also be presented, with a caution statement, where appropriate, about the inherent risks of unplanned subgroup analyses. The summary statement should report the findings of analyses that had a reasonable ability to detect subgroup differences and should note when analyses were not useful because of inadequate sample size. Examples of appropriate summary statements include:

- The database was not large enough to assess whether there were differences in effects in age, gender, or race subgroups.

- Examination of age and gender subgroups did not identify differences in response to (study drug) among these subgroups. There were too few black patients to adequately assess differences in effects in that population.
- Examination of age and gender subgroups suggested a larger treatment effect in women (possibly resulting from the larger mg/kg dose in women), but no age-related differences. There were too few black patients to adequately assess differences in effects in that population.

D. Presenting Data for Different Types of Outcomes

Data on outcomes of treatment should be presented only if the outcome is of clinical significance.

1. Categorical Outcomes (e.g., success or failure)

For categorical outcomes, the number (or percentage) of outcomes for all randomized patients should be shown. For example, the total sample size for the treatment group, the number of successes, the number of failures, and the number of unknown status should be given. Where informative, those patients whose outcome status is unknown can be further differentiated by including the number who dropped out due to adverse events, the number who were lost to follow-up, or any other pertinent distinction. If only percentages are reported, the denominator

should be included.

2. Continuous Variables

For continuous variables, means or medians, accompanied by the standard deviation, are the usual methods for presenting data. When means or medians are used, the magnitude of variability in patient responses should be discussed, and the number of subjects remaining on study at each time point should be given. When means or medians do not adequately convey the variability of responses, it might be useful to display individual responses (e.g., by graphical representation of the cumulative distribution of responses — see appendix). It is important to include the baseline value when reporting any change (either numerical or percent change) from that baseline.

3. Time-to-Event Endpoints

When time-to-event endpoints (e.g., mortality) are used, median or mean survival alone is not usually an adequate descriptor. Survival curves (or event-free survival curves) and hazard ratios are often effective ways to display such data. Data can also be summarized at specific times (e.g., prevalence at 3, 6, 9, 12 months) or at specific event frequency (e.g., time to 25%, 50%, and 75% prevalence of events). The number of patients evaluated at a given interval or frequency should be specified.

4. Graphs or Tables

Ordinarily a graph or table is more effective than text alone in communicating study results, and one or the other should be used when presenting study results in the CLINICAL STUDIES section. See the appendix for guidance on the use of graphs and tables in the CLINICAL STUDIES section of labeling.

E. Advertising and Promotional Considerations

Advertising and promotional materials make frequent use of statements or data appearing in the CLINICAL STUDIES section. Therefore, the CLINICAL STUDIES section should be carefully scrutinized to ensure that its content does not suggest or imply claims for indications, doses, regimens, or comparative effectiveness that are not adequately supported. Words or phrases that lack a commonly understood meaning (e.g., imprecise quantitative terms), are not easily defined, are vague, are misleading, or are promotional in tone should be avoided. Examples include large or small (instead use actual size or amount), well-conducted (instead provide specifics about the study design), extensively studied (instead provide specifics about the database), rapid (instead specify change/unit time), trend (instead provide specifics about the outcome), potent (instead give the size of the effect), pivotal study (instead describe as major efficacy study), and highly significant (instead provide the confidence interval).

F. Updating the Clinical Studies Section

The CLINICAL STUDIES section should be updated when new, important information becomes available. Outdated information should be promptly revised or replaced.

APPENDIX

Presenting Study Results in Tables and Graphs

I. INTRODUCTION

This appendix provides guidance on the use of graphs and tables in the CLINICAL STUDIES section of labeling. When clinical data are to be presented in some detail, ordinarily tables and graphs are better than text alone because they convey the desired information more effectively. The following general principles apply to the use of tables and graphs:

- Tables and graphs should depict study results clearly, fairly, and accurately.
- Text accompanying a table or graph should avoid needless repetition of information that would be clear from viewing the table or graph. The text should serve as an aid to interpreting the most important data presented in the table or graph. Often the statement, "Results are found in Table X," is sufficient.
- Tables and graphs should be next to the text that mentions the table or graph, and that text should refer to the table or graph. Small tables can be embedded in the text.
- Tables and graphs should have clear titles and clearly labeled axes to limit the need to use text to explain what is portrayed.

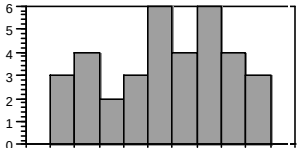
II. GRAPHS

A. Use of Graphs

- To present a large amount of data, such as individual patient data points (cumulative responses).
- To illustrate changes over time.
- To illustrate differences in magnitude of response, particularly where more than two treatment groups are being compared.
- To convey dose-response information.

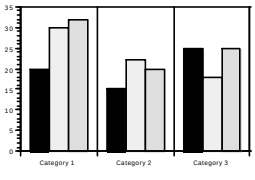
B. Graphs Most Commonly Used in the Description of Clinical Trial Efficacy Data

Histogram



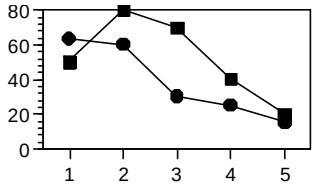
Illustrates individual patient data by presenting the number or percentage of patients (y-axis) exhibiting a given response (x-axis) over the whole response range.

Bar Graph



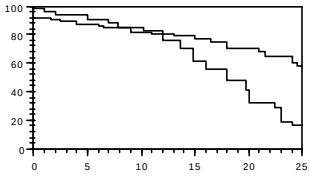
(1) Compares subgroups where each bar represents a subgroup within a treatment arm, and the length of the bar represents the group response for the outcome variable, (2) shows the percentage or frequency of patients (the y-axis) exhibiting a categorical response, and (3) displays the principal results of several similar trials. In most cases, it is helpful to include error bars. A bar graph should not be used to illustrate just a few numbers that could be summarized better in a table. 3-D graphs should be avoided because they make comparisons between bars very difficult. Stippling or other small patterns in bars should also be avoided because they can be difficult to see after reduction or reproduction.

Line Graph



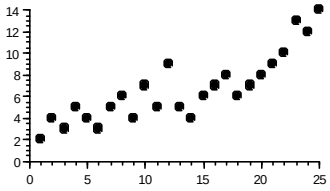
Most often illustrates responses (y-axis) over time (x-axis). It is helpful in many cases to include error bars and number of patients remaining on study treatment at each time point. Similar graphs can be used to show dose response with response on the y-axis and dose on the x-axis.

Survival Curve



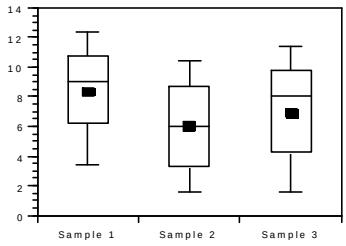
Depicts time-to-event data for events like death or recurrence of disease. Usually, Kaplan-Meier estimates of the proportion of patients surviving are plotted, but some plots show the raw cumulative incidence rates over time.

Scatter Plot



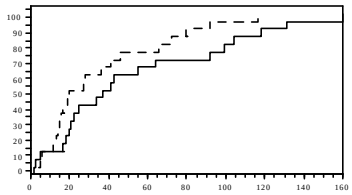
Shows the relationship between two (usually continuous) variables for individual patients, such as response (y-axis) related to blood levels or some other measure (x-axis).

Boxplot



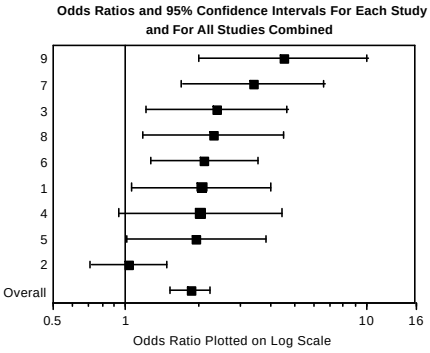
Illustrates the distribution of data for a single group. Several plots in a single graph are useful for comparing distributions. Boxes represent the range of values from the 25th percentile to the 75th percentile. The definition for the length of the whiskers (lines extending out from each end of the box) varies with software packages and should be defined with the plot (e.g., the ends represent the 10th and 90th percentile, the minimum and the maximum, or lower adjacent value computed as 1.5 times the interquartile range minus the 25th percentile and the upper adjacent value computed as 1.5 times the interquartile range plus the 75th percentile).

Cumulative Distribution Plot



Depicts individual patient data. The y-axis represents a cumulative percentage of subjects with a response at least as large as the effect shown on the x-axis. These distributions can be graphed using bars or using a step graph. A cumulative distribution plot could call for additional text that describes how to read the graph. For example, the following text could be helpful: "A curve shifted to the left represents a more effective treatment. A response of at least *x* was seen by *a percent* of the patients on New Drug and *b percent* of the patients on Placebo."

Meta-analytic Graph



Depicts summary estimates (usually a treatment difference) for several studies (or centers) on one graph. It is useful for illustrating a lack of consistency across studies. Ordering the responses by magnitude enhances the visualization of the effects. Similar displays can show results in demographic or other subsets (e.g., disease severity, background therapy).

C. Features of a Good Graph

- **Title:** The following information should be included in the title: the name of the study, the type of data, the timepoint, and important features of the patient population (e.g., Intent-to-Treat, Evaluable, age range if relevant). For ease of reading, the beginning of each word should be capitalized, not every letter in the title.

- **Axis Label/Title:** Ordinarily each axis should be labeled. Units of measurement should be included.

- **Ticks and Grids:** Ticks for each axis should be labeled so that the reader does not have to interpolate to understand the data. A graph will appear less cluttered if ticks face away from the graph and if grids are eliminated.

- **Axis Scale:** The treatment effect should not be exaggerated (e.g., interruptions) by the scale of measurement (generally the y-axis), but the scale of measurement should show the scale of the efficacy variable. The scales should be consistent for like graphs within the label. Differences in scales among labels of drugs in the same class should be avoided if possible, as they can lead to misleading comparisons.

- **Symbols:** Symbols should be easily distinguishable by size, shape, or fill (e.g., open symbols for placebo and closed symbols for treatment).

- **Footnotes:** A footnote should be used if further information would be helpful to explain the content of the graph (e.g., the meaning of a term used, the meaning of a symbol). Ordinarily, statements interpreting the graph should be included in the text accompanying the graph and not in a footnote.

- **Error Bars:** It should be clear from the graph which measure of variability is used to define the error bars (standard deviation, standard error, percentiles).

- **Font Size:** To ensure the graph is readable, careful thought should be given to selecting the font size for labels and symbols. Graphs in labeling and in the *Physicians Desk Reference* (PDR) are often reduced in size to about 60 mm by 50 mm.

- **Legend:** The graph should not be overpowered by the legend. Labels directly on the graph are preferable to a legend.

- **Sample Size:** Including sample sizes for each group often helps the reader interpret the graph. Sample sizes can be identified in text within the graph or in a small table just below the graph.

- **Uncertainty of Treatment Effect:** Differences should be accompanied by the appropriate measure of uncertainty (confidence interval or p-value). Differences that are not statistically significant should be identified as such.

III. TABLES

A. Use of Tables

- To present simple, descriptive statistics such as medians, means, standard deviations, and sample size for both treatment groups or for only a few time points.

- To summarize data from more than one efficacy variable.

- To present exact values if that information is desirable.

B. Features of a Good Table

- **Title:** The following information should be included in the title: the name of the study, the type of data, the timepoint, and the patient population (Intent-to-Treat, Evaluable). For ease of reading, the beginning of each word should be capitalized, not every letter in the title.

- **Units:** The units of measurement for the data presented should be included in the table, either in the title or column headings. When presenting percentages, it is helpful to include the percent sign, particularly when several numbers are included on one line (such as mean percentages and

sample sizes). Only include the number of digits after the decimal that are significant or meaningful.

- **Sample Size:** The sample size for each treatment group should be included in the table. Also, the age range should be identified when relevant.

- **Baseline Data:** Baseline data should be included whenever applicable.