POLICY

OFFICE OF TRANSLATIONAL SCIENCES

Good Review Practice: Statistical Review Template

Table of Contents

PURPOSE

This MAPP establishes procedures for documenting the review of original New Drug Applications (NDAs) and Biologics License Applications (BLAs), NDA/BLA amendments in response to action letter, and efficacy supplements using good review practices in the Office of Biostatistics (OB), Office of Translational Sciences (OTS), Center for Drug Evaluation and Research (CDER).

POLICY

1. All reviewers within the Office of Biostatistics will use the Statistical Review Template (Attachment 1).
2. The Statistical Review Template will be used to document primary statistical reviews of all original NDAs and BLAs, and NDA/BLA amendments in response to action letter and efficacy supplements.
3. Conventions of CDER's Style Guide should be followed in completing the statistical review.

RESPONSIBILITIES

- The statistical reviewer will complete each review of designated submissions using the statistical review template. The statistical reviewer will engage in scientific and regulatory dialogue concerning his or her analyses and conclusions, and share a draft review, with the

statistical team leader and other statistical supervisors to develop complete and scientifically valid review perspectives. However, the final conclusions and recommendations in the statistical review should reflect the statistical reviewer's own opinion and should emphasize that the conclusions and recommendations are based on the review of the relevant portion of the application, not the entire application.

- The statistical team leader will promote consistent use of the statistical review template by statistical reviewers. The statistical team leader will engage each assigned statistical reviewer in scientific and regulatory exchanges regarding reviews before finalization of the statistical review. When the statistical reviewer's conclusions and/or recommendations differ from those of the statistical team leader, the statistical team leader should encourage the statistical reviewer to document his or her own conclusions and recommendations in the statistical review. In such cases, the statistical team leader is expected to write his or her own review, noting the reasons for any differences in conclusions and recommendations from those of the statistical reviewer.

- Division and office directors will promote consistent use of the statistical review template, provide scientific and regulatory perspective on review issues, and encourage statistical reviewers to document in the statistical review their rationale for their own perspectives.

PROCEDURES

To document statistical reviews, statistical reviewers will use the statistical review template by following the instructions in the attachments to this MAPP. The template is annotated to provide additional explanations of the content for each heading and subheading.

REFERENCES

1. FDA/CDER, 2012, CDER Style Guide, Style and Formatting for CDER Documents.
2. FDA, 2010, Center for Drug Evaluation and Research, MAPP 6010.3 Rev. 1: Good Review Practice: Clinical Review Template.

SUMMARY OF CHANGES

1. MAPP 4000.8 NDA and MAPP 4000.8 BLA have been consolidated into a single MAPP.
2. Changes to the organizational structure have been incorporated.
3. Revised template appears as Attachment 1.

EFFECTIVE DATE

This MAPP is effective upon date of publication.

CHANGE CONTROL TABLE

Effective Date	Revision Number	Revisions
4/20/05	Initial(s)	The Biostatistics Biologics Licensing Application Review template policies were posted as MAPP 4000.8 BLA. Also, the biostatistics New Drug Application Review template was posted as MAPP 4000.8 NDA. The Template for both MAPPs was posted as MAPP 6010.3
7/30/12	Rev. 1	A single MAPP, issued as CDER MAPP ####.#, is issued. The new Directive supersedes both MAPP 4000.8 postings. Changes to the organizational structure is incorporated, and the revised template it attached as Attachment I.

ATTACHMENT 1 – Statistical Review and Evaluation

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA 123-456
Supplement #:	As applicable. This # can be found in DARRTS under "Submission Type-Number". If the submission is listed as "Original", then it should be omitted.
Drug Name:	Trade name (generic name) strength and dosage form
Indication(s):	List indication(s) of treatment under review
Applicant:	Identify applicant of the submission
Date(s):	At a minimum, the date submitted. Other dates, such as date received by CDER or reviewer, PDUFA due date, or review completion date, may be given.
Review Priority:	Priority, Standard, etc.
Biometrics Division:	
Statistical Reviewer:	
Concurring Reviewers:	Statisticians who reviewed and signed this review (e.g., statistical team leader).
Medical Division:	The medical division to which this NDA/BLA is assigned.
Clinical Team:	At a minimum, the medical officer(s) reviewing this application. The medical team leader and medical division director may be listed as well.
Project Manager:	

Keywords:

Link to keywords:

http://intranetapps.fda.gov/scripts/ob_apps/ob/eWork/uploads/eWork/2009/Keywords-in-DFS.htm

Table of Contents

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The Executive Summary should include a “bottom-line” statement without equivocation and should be written in plain language appropriate for educated lay as well as technical audiences. It should be a concise summation that adequately relays relevant information regarding the submission (i.e., not a cut and paste of section 5). Typically, it should not exceed 2 pages. Further details of the evaluation should be documented in the body of the review.

The summary should describe the conclusions of the statistical reviewer's evaluation. The conclusions should be the synthesis of evaluations of all studies under review. Recommendations should be stated within the context of strength of statistical evidence and key findings that may or may not support the claims of the applicant.

This section should also describe key statistical issues and findings that impact demonstration of efficacy/safety. It should summarize and discuss the reviewer's analyses, the extent of evidence in support of claims, statistical issues that may affect the conclusion on efficacy/safety, and any related comments. It should be based on each study reviewed as well as on the collective evidence. In addition to the primary endpoint analysis, the reviewer may also address secondary or subgroup analyses if these are deemed important.

Where necessary, it is recommended that the reviewer provide references to text, tables, and graphs to which readers can refer within the review.

INTRODUCTION

This section will give some information on the drug development for this submission, the studies submitted, and those selected for the review.

Overview

The overview sub-section should include background information regarding the investigational drug, the drug class, and its intended indication. Additionally, this sub-section should include important clinical program milestones reached during the course of drug development which are relevant to the statistical review. In particular, it should include the advice given to the applicant during IND development, whether at meetings or through correspondence.

All relevant studies in the clinical program should be listed. The reviewer should identify those studies selected for full statistical review and evaluation, and briefly explain why they were chosen.

The selected studies should be specifically highlighted with information such as study identification number/name; study type/design; number of treatment arms; indications; number of patients in each arm; location of investigational centers; and proportion of patients enrolled in domestic versus foreign investigational centers. Key information should be presented in a table, such as the example below:

Table: List of all studies included in analysis

	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
Applicant defined study number	*Phase 2/3*				*e.g., critical disease or patient characteristics*

The reviewer may document the overview with sub-sub sections numbered as 2.1.1, 2.1.2, 2.1.3, etc. Examples of these are: Class and Indication; History of Drug Development; Specific Studies Reviewed; and Major Statistical Issues.

Data Sources

Data sources include all material reviewed, e.g. applicant study reports, data sets analyzed, and literature referenced.

If data were provided electronically, the location/names of data sets should be documented. Indicate which data formats were used (e.g., SDTM) and whether software code was submitted. The full electronic path of these data should be specified according to the CDER Electronic Document Room (EDR) naming convention. Evaluations based on literature reviews should identify the source and extent of available raw data. If only a paper version is submitted, then references to volume, section, page, table, and/or graph should be specified. The quality and integrity of the data will be discussed in Section 3.1.

STATISTICAL EVALUATION

This section will present the detailed review of selected studies to be reviewed from the NDA and possibly from other sources such as published literature.

Typically, an NDA submission has one or more studies for a single indication. However, some submissions involve multiple indications, multiple special populations, or multiple disease etiologies. Therefore, the reviewer may organize the review by individual study, by indication, by special population, or by etiology of disease.

Each study should be reviewed in the main body of the text. Mathematical details and derivations can be put in the Appendix.

Under the main heading of Section 3 or under sub-section headings 3.1 and/or 3.2, the reviewer may describe the organization of sub-sub-sections. For example, an ordering of sub-sub-sections by indication may be described here.

Data and Analysis Quality

Review the quality and integrity of the submitted data. Examples of relevant issues include the following:

- Whether it is possible to reproduce the primary analysis dataset, and in particular the primary endpoint, from the original data source
- Whether it is possible to verify the randomized treatment assignments
- Findings from the Division of Scientific Investigation or other source(s) that question the usability of the data
- Whether the applicant submitted documentation of data quality control/assurance procedures (see ICH E3,[1] section 9.6; also ICH E6,[2] section 5.1)
- Whether the blinding/unblinding procedures were well documented (see ICH E3, section 9.4.6)

[1] http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073113.pdf

[2] http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073122.pdf

- Whether a final statistical analysis plan (SAP) was submitted and relevant analysis decisions (e.g., pooling of sites, analysis population membership, etc.) were made prior to unblinding.

Reviewers may also want to *consult Guidance for Industry – Computerized Systems Used in Clinical Investigations* (2007)[3] to get an idea of the regulatory expectations for clinical data systems.

Applicants are expected to submit data of high quality and make it possible for the FDA to reproduce their results. In turn, FDA reviewers should provide adequate documentation so that the applicant or another data user could reproduce their independent findings. The level of documentation needed will depend on the complexity and novelty of the analysis. If an ordinary ANOVA or ANCOVA is used, for example, it would suffice to identify the dependent and independent variables. If a more unusual analysis is performed, then it may be necessary to provide code. The code should be either included in the report or put in an appropriate digital archive.

Describe the level of effort needed to process the data. Any collaboration with the Computational Science Center should also be described in this section.

Evaluation of Efficacy

This section may be omitted if the review is focused entirely on safety.

The assessment of efficacy for each study reviewed should include a description of the study design; primary and secondary efficacy endpoints; demographic and baseline characteristics; patient disposition; statistical methodology used; applicant's results; and the reviewer's findings.

In addition to the registration studies, the reviewer should also consider the results of other relevant studies (positive or negative). These studies should be discussed in this section and considered when determining the overall strength of evidence regarding efficacy in Section 5.1.

The format of this section will depend on the application being reviewed. Reviewers are encouraged to organize the review by adding sub-sub sections (3.2.1, 3.2.2, etc.), examples of which are described as follows.

Study Design and Endpoints

A description of the design, endpoints, and conduct of each relevant study should be included in the review. The reviewer should identify the aspects of the design that may be inappropriate or introduce bias to the final results. If the study is adaptive, uses an

[3] http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070266.pdf

enrichment design, or is designed to show non-inferiority to an active comparator, then these features should be clearly described. If there are multiple endpoints, then any planned approach to control the Type I error rate should also be specified here.

The reviewer may discuss the relevance or appropriateness of the design and endpoints. The review may also include a brief description of the current thinking of the reviewing medical division regarding the appropriateness of the design and endpoints. If a non-inferiority design or adaptive design is used, then detailed description of relevant information should be documented. The quality of the endpoint assessment, including any adjudication procedures that were used, may be discussed in this section.

Statistical Methodologies

The reviewer should describe the statistical methodologies used by the applicant, as well as the reviewer's alternative methodologies (with explanation). The reviewer may document technical discussion, mathematical derivation, or presentation of formulae in an appendix rather than in the main body of the review.

In particular, the reviewer should describe the Applicant's pre-specified methods for handling missing data, including any sensitivity analyses. The reviewer's own analyses should also be described.

The appropriateness of adjustment for stratification factors and/or covariates in the analysis and whether or not such adjustments were pre-specified may be discussed in this section. If the reviewer's analysis incorporates a different set of stratification factors or covariates in order to examine the robustness of the Applicant's findings, those analyses should be presented here.

If meta-analyses were submitted, then the methodology for selecting and analyzing the studies should also be presented.

Patient Disposition, Demographic and Baseline Characteristics

A description of the patient populations should be included in the review. This can be done by including tabulations of percent dropouts, reasons for dropouts, and protocol violations. One may also consider presenting the time to dropout graphically, e.g., using Kaplan-Meier curves. Describe the pattern of treatment and study discontinuation and its impact on missing data. For example, indicate whether patients who discontinued treatment were followed-up for efficacy and safety (retrieved dropouts). A table showing the demographics and baseline characteristics of the subjects in each treatment arm should be included. Noteworthy differences between the study population and the intended treatment population may be noted, particularly if there is a concern that these difference(s) undermine the generalizability of the results. Additionally, the reviewer must describe the primary efficacy analysis population. Other analysis sets may also be included.

Results and Conclusions

The reviewer should provide a brief summary of the applicant's results and conclusions. For each study examined, the statistical reviewer should discuss efficacy findings, the strength of the evidence, and statistical issues. The reviewer must provide tables and/or graphs to describe the extent to which study results support the efficacy claim.

Detailed discussion may not be necessary for those issues for which both the applicant and the reviewer generally concur. However, any differences between the conclusions of reviewer and the applicant need to be explained. If a table, figure, or text provided by the applicant is included in the review, then provide an appropriate reference, e.g., "Table 16 in the Clinical Study Report for Study 3". It may also be appropriate to refer to the module and/or page numbers. In some circumstances, it may be helpful to also explicitly identify tables or figures that were produced by the reviewer.

During the course of the review, the reviewer should address various statistical issues. Specific issues that may be discussed in this sub-section include, but are not limited to, the influence of covariates, missing data imputation methods, sensitivity analyses, adaptive design, enrichment design, non-inferiority, and subgroup analyses. The statistical issues, their resolution, and any impact on efficacy assessment need to be discussed.

Evaluation of Safety

This section is devoted to safety analyses that are crucial for drug approval and labeling, even when the main focus of the review is on efficacy. Note that this section (or subsections provided below) may not be applicable to all reviews; therefore, the scope of this section should be discussed with the clinical review team.

If the Applicant submitted a pre-specified safety analysis plan, the reviewer should describe how well they followed the plan and discuss any deviations from it in this section. Additionally, the reviewer should comment on whether there was a data monitoring committee and any stopping rules for safety. When applicable, the reviewer should provide background about known (or anticipated) safety concerns of the drug class. These safety concerns should be distinguished from those that are expected because of the study population or underlying disease, but not necessarily known to be associated with the drug or drug class. The reviewer may reference guidances pertaining to the particular drug class as needed.

The reviewers are recommended to use subsections similar to those specified below, especially for focused safety reviews, in which the studies in the application were predefined to evaluate safety.

3.3.1 Safety Analysis Population(s) and Endpoint(s)

The safety analysis population definition should be clear and indicate whether patients were analyzed as randomized or as treated. For example, the safety analysis population might comprise patients who received a minimum number of doses of treatment or patients who

underwent treatment for a specified duration. If the reviewer uses a safety analysis population that is different from that used by the Applicant, the reviewer should provide rationale.

In addition to the safety analysis population, the reviewer should provide definitions of the safety outcome(s) or endpoint(s) being analyzed. The reviewer should comment on whether outcomes (or endpoints) were pre-specified in the study protocol, and whether they were systematically collected or spontaneously reported. Also, the reviewer should note whether endpoint definitions were objective (e.g. a known laboratory test for disease presence) or subjective (e.g. an arbitrary cut-off for "high risk"). Subjective endpoints may introduce bias or obscure true safety issues. The relevance of the endpoints and measurements should be discussed with the clinical team. Additionally, the reviewer should address how these outcomes were reported (for example, MedDRA version 13.0) and whether they were adjudicated. The reviewer should discuss any study design issues that limited ascertainment of the safety outcome(s), such as when the study stops after the patient meets the efficacy endpoint but before a safety outcome has occurred. The reviewer should summarize all studies (e.g. pre-clinical, Phase I, etc.) considered for safety and discuss whether the outcomes were defined consistently across these studies.

3.3.2 Data Quality

Any problems, inconsistencies, or incompleteness of the datasets that prevent safety assessments should be noted in this section, including evidence of reconciliation between adverse event reports and reports of study discontinuations due to safety. The reviewer should summarize the amount of missing data and whether the missingness appears to be non-random; non-random missingness can bias safety results, particularly if the missingness is in variables that are deemed clinically important.

3.3.3 Statistical Methods

The reviewer should describe and comment on the appropriateness of the Applicant's safety analyses, particularly if suitable for rare events, and if the analyses have adequate focus on safety concerns known a priori for the drug or drug class. It is also important that the reviewer assesses whether assumptions for the analysis methods are met. The reviewer should discuss and provide justification of any different or additional analyses from those done by the Applicant. The reviewer might consider sensitivity analyses to assess the impact of missing data. Additionally, the reviewer should address if the safety outcomes that occurred in the open-label or extension phases of trials were included in the analyses.

For focused safety studies, the reviewer should discuss whether the Applicant's proposed sample size was sufficient to rule out a desired risk margin with adequate power.

<u>Integrated Analyses</u>

The reviewer should assess whether the characteristics of the studies are similar enough to justify integrating. Some cases where integrating across studies might be problematic include

differences in the frequency or definitions of the safety outcomes, inclusion/exclusion criteria, randomization ratios, comparator arms, and discontinuation proportions.

3.3.4 Results and Conclusions

The reviewer should describe the drug (and placebo/control) exposure time and follow-up time, potentially by calculating the mean or categorizing the exposure time as determined by the clinical team, and address whether the exposure or follow-up times were sufficient to adequately capture the safety outcomes of interest. For example, if there is a concern of carcinogenicity, much longer follow-up may be required than for an allergic reaction. The reviewer should also evaluate the relationship between the safety outcome and withdrawal; distinction between withdrawal from trial and withdrawal from treatment is recommended. The reviewer should note whether there is a pattern for censoring/dropout, especially if the censoring/dropout rates between the treatment and control arms differ.

The reviewer should summarize adverse events that occurred at a high frequency and those that are of interest to the clinical team. Where applicable, the exposure-adjusted risk of events and the corresponding confidence intervals should be presented. The reviewer might also consider reporting safety outcomes grouped according to MedDRA hierarchy, severity of the event, or other categories recommended by the clinical team.

For safety outcomes that were not pre-specified, a description of the size and direction (negative or positive) of the risk estimates and confidence intervals may be sufficient, rather than conclusions of statistical significance based on p-values. The reviewer should also be cautious of interpreting statistical significance from subgroup analyses and crude pooled data because the interpretation of results can be misleading (e.g. Simpson's Paradox). In addition, statistically non-significant findings may not imply the product is safe. Thus, the reviewer should carefully assess any claims made by the Applicant in the study report or label about product safety from such findings. For composite safety outcomes, the reviewer should provide risk estimates for each component of the composite and discuss whether the results are highly influenced by a particular component.

For reviews involving meta-analytic methods, reviewers are encouraged to summarize, at minimum, the quality of the studies included (e.g. if there is differential dropout, high percentage of missing data), how heterogeneity across studies was assessed (including clinical and statistical assessments), and the final statistical model (random effects or fixed effect) used to combine studies along with reason(s) for model choice. Additionally, when applicable, the reviewer should specify the inclusion and exclusion criteria used to select studies to be included in the meta-analysis. It is worth examining whether one study appears to be driving the overall results. If so, the reviewer should consider whether the population included in that study was the same as the other studies, or whether there was some greater predisposition to risk that may be present in that trial.

Where applicable, presenting safety findings in tables and graphs are encouraged, refer to the appendix of "Guidance for Industry: Clinical Studies Section of Labeling for Human Prescription Drug and Biological Products-Content and Format."

In this sub-section, the reviewer should discuss any significant limitations that impact the statistical safety evaluation. The reviewer should also provide recommendations for the label or for further studies (e.g. post-market requirements) to address the safety concerns identified in the review.

In addition to the topics presented above, if the focus of the review is safety, then sub-sections similar to 3.2.1-3.2.4 should be incorporated.

Benefit-Risk Assessment (Optional)

The purpose of this optional section is to address benefit-risk issues in the application, i.e., if the benefit is large enough to justify the risk of the treatment. As with the previous sub-section, the scope of this section should be determined in coordination with other relevant reviewers. Issues that may be considered in this section include, but are not limited to, the impact of withdrawals on benefit-risk and the association between adverse events and the key efficacy outcome(s). Number needed to treat (NNT) and/or number needed to harm (NNH) may be included, provided that they can be estimated with sufficient precision.

FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The reviewer should describe efficacy (safety) results across subgroups defined by gender, race, age, and geographic region. Other subgroups such as those based on baseline characteristics may be included depending on their relevance, representation in the clinical studies, or on the disease being reviewed.

Under the main section heading 4, the reviewer may give an overview of which subgroups are relevant to the application under review.

Gender, Race, Age, and Geographic Region

In this sub-section, the reviewer should describe efficacy (safety) results across subgroups defined by gender, race, age (for example less than 65 versus greater than or equal to 65 years), and geographic region (for example, US vs. non-US).

The reviewer should include descriptive statistics for the defined subgroups. Inferential statistics such as the results of tests for treatment by subgroup interactions may also be included. Significant interaction test results should be fully explained, for example by including graphics depicting the results. The reviewer should exercise caution when synthesizing the data across studies.
Scientifically valid methods should be employed when drawing inferences from pooled data.

The impact of a subgroup difference may be briefly addressed here and more fully explained in Section 5.1 or *vice versa*.

Mention should be made if no conclusions can be drawn due to lack of representation, limited sample size, etc. If, for example, the studies were conducted in one gender only, a brief statement should indicate that gender analysis was not applicable.

Other Special/Subgroup Populations

Other subgroups could be defined by baseline characteristics. These should be included depending on their relevance, on their representation in the clinical studies, or on the disease being reviewed.

If no other subgroups other than those in sub-section 4.1 are reviewed, the reviewer should indicate here that "No other subgroups were analyzed".

SUMMARY AND CONCLUSIONS

Statistical Issues

Only the statistical issues that impact the overall conclusions should be described here. The main statistical issues should be summarized study by study, as well as collectively, for all studies in the review. It may be necessary to refer to other sections of the review or to appendices to provide sufficient detail. Resolution of these issues and their impact, if any, on overall efficacy assessment need to be briefly discussed.

Discussion of the statistical issues serves to justify the comments and conclusions of the reviewer and brings attention to these issues for future trials.

Examples of important statistical issues that may affect the results are the following:

- Breaking the blind
- Unblinded or unplanned interim analyses
- High percentage of dropouts
- Inappropriate imputation for missing values
- Change of primary endpoint during conduct of the trial
- Dropping/adding treatment arms
- Sample size modification
- Inconsistency of results across subgroups
- Type I error inflation due to multiplicity
- Planned and unplanned adaptations
- Non-Inferiority

Where necessary the reviewer should provide easy-to-read yet fully informative tables, graphs, and text to collectively describe the overall extent to which study results support the efficacy claim.

Collective Evidence

The reviewer's summary of the collective evidence of effectiveness and/or safety is a compilation of the main findings from all studies reviewed. This summary may include treatment estimates obtained by combining studies where appropriate. As was advised for special/subgroup populations, the reviewer should exercise caution when pooling data across studies. Scientifically valid methods should be employed when drawing inferences from pooled data. Well-controlled studies that do not demonstrate drug effect should be considered when determining the strength of evidence.

Conclusions and Recommendations

The statistical reviewer's conclusions should be based on collective evidence provided by the entire application, as described in Section 5.1. They should be made within the context of whether the statistical results do or do not provide adequate evidence to support the claims proposed in the NDA. If the reviewer's conclusion differs from that of the applicant, the reviewer needs to briefly mention the reasons for these differences.

Labeling Recommendations (as applicable)

This section should address any major areas of disagreement with the applicant's proposed label, for example not to include subgroup or secondary endpoint analyses without multiplicity adjustments. It may be amended after the label is finalized.

APPENDICES (Add When Needed)