Clinical Epidemiology and Biostatistics Unit (CEBU)
STANDARD OPERATING PROCEDURE (SOP)

murdoch children's research institute **Version: 1.1**	**Title:** **Statistical Analysis Plan (SAP) for Final Analyses**

Murdoch Children's Research Institute

Prepared by: Katherine Lee and Francesca Orsini

Date: 20th October 2011

Authorised by: Katherine Lee

Date: 19 April 2021

Change History

20th October 2011: Initial version of guideline finalised – Version 1.0

16 April 2021: Initial version reviewed, cosmetic changes made – Version 1.1

Contents

1. PURPOSE

To describe the process for developing, reviewing and approving the Statistical Analysis Plan (SAP) for the final analysis of a study.

2. APPLICABILITY

This procedure applies to all studies for which the SAP development is delegated to CEBU. The section relevant to the unblinding process applies only to blinded studies.

3. ROLES AND RESPONSIBILITIES

Statistician: responsible for the development of a detailed document regarding the methods of statistical data analysis proposed for the study.

Review team: responsible for the approval of the SAP. This team could include the principal investigator, the staff involved in the study conduct and the Sponsor's reference person (in the case of commercially sponsored studies).

4. DEFINITIONS

Database lock: once data management activities (data cleaning, queries resolution, adverse event reconciliation and coding activities) have been performed and there are no pending issues the database is considered locked and it cannot be modified.

Unblinding: when a member of the study team requests to know the treatment allocation of one more participant in the study.

Interim analysis: any analysis intended to compare treatment arms with respect to efficacy or safety at any time prior to the formal completion of a study.

5. INTRODUCTION

A SAP is a document that contains a technical and detailed description of the statistical analysis required for the study described in the protocol. It is completed after finalising the protocol.

The plan should be reviewed and possibly updated as a result of the blind review of the data and should be finalised before breaking the blind. Ideally a blind review of the data is a pre-analysis review, blinded to treatment, where decisions that may lead to changes to the analysis that was foreseen in the protocol (exclusion of participants or data from the analysis sets, outlier definitions, important covariates identified in other recent research added to the model, etc...) are made.

The SAP must be finalised and approved before any analysis is carried out. Formal records of when the SAP is finalised as well as when the blind is subsequently broken should be kept in the electronic study folder within the study directory and the paper copy in the study binder.

Guidance for the development of a SAP is provided in ICH E9 "*Statistical Principles for clinical trials*" and *ICH Point to consider* documents. The contents of the SAP should clarify all the items relevant for the analysis of the study data, in particular providing explicit details of the statistical method to be used, elaborating on the details in the protocol.

6. PROCEDURE

Ideally the SAP should be developed using an appropriate template (SAP template final analysis), including all other relevant information raised during the study conduct and blind review.

The SAP should contain:

- The study objectives and the key parameter to be estimated or hypothesis to be tested
- The study design, the randomisation process and sample size calculation as per study protocol
- The inclusion and exclusion criteria
- The definition and the role of the analysis populations (per protocol, intention-to-treat, safety, and other population if required), specifying the main statements for the patients' exclusion from the populations
- A detailed description of general methodology for summarising demographic/baseline characteristics, efficacy and safety variables (where applicable)
- The method for analysis of primary and secondary outcome variables
- Handling of missing values, dropout
- Any planned data transformation
- Methods of exploratory analysis, if any

Optionally, a list of Tables and a list of Patient Data Listings to be included in the statistical report (see SAP template final analysis).

Once drafted, the first version of the SAP should be sent, together with lists of tables and listings (if required), to the assigned SAP review team. The review team reviews the first version of the SAP documentation and returns any comment to study statistician. The statistician addresses the review comments and discusses any conflicting and outstanding issues with the SAP review team.

Once the review comments have been incorporated and all the outstanding issued have been addressed, the SAP documentation should be sent to the review team for final approval. Approval via e-mail is acceptable.

If the planned analysis needs to be changed after the SAP has been approved, the statistician should make the necessary changes, then submits the updated version of the SAP to the review team/SAP approver, and the approval process is repeated as described above. When changes to the planned analysis are made after unblinding, these changes should be documented in the statistical and clinical study report and not the SAP.

GENERAL PRINCIPLES

Study objectives and the key parameters to be estimated and/or hypotheses to be tested

In line with the protocol, the SAP should precisely specify the parameters that are to be estimated and/or hypotheses to be tested including clear definition of the outcomes that are to be used in order to satisfy the primary and secondary objectives of the study. The primary outcome variable (primary endpoint) should be the variable capable of providing the most clinically relevant and convincing evidence directly related to the primary objective of the study.

Study design, randomisation process and sample size calculation

The study design, the randomisation process and the sample size estimation should be clearly specified. This should come directly from the study protocol.

Analysis populations, protocol violations and Inclusion/Exclusion criteria

If all participants recruited into a study satisfied all entry criteria, followed all study procedures perfectly with no losses to follow-up, and provided complete data records, then the set of participants to be included in the analysis would be self-evident. The design and conduct of a study should aim to approach this ideal as closely as possible, but, in practice, it is doubtful if it can ever be fully achieved. Hence, the SAP should address anticipated problems in terms of how these affect the analysis of interest. The anticipated irregularities in study conduct might impair a satisfactory analysis, including various types of protocol violations, withdrawals and missing values. Hence, the SAP should consider ways to handle the problems that do occur in the study in the analysis of data, taking into account the following main issues:

- *Analysis populations.* The set of participants whose data are to be included in the main analyses should be defined (where the main analysis is defined as the key results from the study in terms of the primary and secondary objectives, as opposed to the sensitivity analysis used to check the robustness of the findings). Decisions concerning the population for analysis, guided by the principles of minimising bias and to avoid inflation of type I error, should be clearly described in the SAP.
-

Populations to be used in the statistical analysis as defined in the protocol (e.g. safety, Intent-To-Treat, Per Protocol) should be clearly defined in accordance with the intention-to-treat/per-protocol principles where required. The precise reasons for excluding participants from the analysis sets should be fully defined and documented in a manner appropriate to the circumstances of the specific study.

- *Protocol violations.* The SAP should specify what defines a protocol violation/deviation and how these will be handled in the analysis. Protocol violations should be summarised in the final study report, with the SAP defining how these should be presented. Relevant protocol violations may include errors in randomisation/treatment assignment, the use of excluded medication, poor compliance to protocol and to treatment, loss to follow-up and missing data.

- *Inclusion/Exclusion criteria*. All the inclusion and exclusion criteria foreseen in the protocol should be listed in a table. If there is reason to believe that there were additional entry criteria, not defined in the protocol, the implications of these should be discussed. For example, some investigators may have excluded patients who were particularly ill or who had particular baseline characteristics. The criteria for exclusion at entry into the study should be specified and the rationale (e.g., safety concerns, administrative reasons or lack of suitability for the study) provided. The predetermined reasons for removing patients from therapy or assessment observation, if any, should be described, as should the nature and duration of any planned follow-up observations in those patients.

Method of analysis for the primary and secondary outcomes

The statistical methods to be used to analyse the study data and to test the study hypotheses should be described for the primary and the secondary outcome variables. Estimates of target parameters should be accompanied by confidence intervals, whenever possible. A description should be given of any intentions to use baseline data to improve precision or to adjust estimates for potential baseline differences or confounding and the statistical method to be used, for example by means of analysis of covariance. It is important to clarify whether one- or two-sided tests of statistical significance will be used, and in particular to justify prospectively the use of one-sided tests if applicable. The particular statistical model chosen should reflect the current state of medical and statistical knowledge about the variables to be analysed as well as the statistical design of the study. All covariates to be included in multivariable analyses should be fully specified. In the choice of statistical methods due attention should be paid to the expected distribution of all outcome variables. When making this choice (for example between parametric and non-parametric methods) it is important to bear in mind the importance of providing statistical estimates of the size of treatment effects together with confidence intervals (in addition to significance tests).

The primary analysis of the primary outcome variable should be clearly distinguished from supporting analyses of the primary or secondary variables.

Handling of dropout, missing values

Missing values represent a potential source of bias in a study. Hence, every effort should be undertaken to ensure that all data are collected where possible. Nonetheless there will almost always be some missing data and it is important to specify how this will be handled in the statistical analysis. The SAP should clearly define how different kinds of missing data will be handled according to the study design, the study objectives affected (primary or secondary) and appropriate assumptions about the reasons for missingness (e.g. missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). No universally applicable methods of handling missing values can be recommended. An investigation should be made concerning the sensitivity of the results of analysis to the method of handling missing values, especially if the number of missing values is substantial: how this sensitivity analysis is going to be carried out should be clearly specified in the SAP.

In some cases it may also be necessary to specify a sensitivity analysis to explore the influence of outliers. Clear identification of a particular value as an outlier is most convincing when justified clinically as well as statistically, and the clinical context will then often define the appropriate action. Any outlier procedure set out in the SAP should be such as not to favour any treatment group a priori. Once again, this aspect of the analysis can be usefully updated during blind review.

Data Transformation

The decision to transform key variables prior to analysis is best made during the design of the study on the basis of clinical relevance and transformation used in similar data from earlier studies. Transformations (e.g. square root, logarithm) should be specified in the SAP and a rationale provided, especially for the primary outcome variable. The general principles guiding the use of transformations to ensure that the assumptions underlying the statistical methods are met are to be found in standard texts; conventions for particular variables have been developed in a number of specific clinical areas. The decision on whether and how to transform a variable may be influenced by the preference for a scale which facilitates clinical interpretation. Subsequent clinical interpretation should be carefully considered and justified in the protocol.

Methods of exploratory analysis

As with all hypothesis driven/confirmatory studies, exploratory studies should have clear and precise objectives. However, in contrast to confirmatory studies, their objectives do not lead to simple tests of pre-defined hypotheses. In addition, exploratory studies may sometimes require a more flexible approach to design so that changes can be made in response to accumulating results. Their analysis primarily involves data exploration; tests of hypothesis may be carried out, but the choice of hypothesis may be data dependent, so only informal interpretation will be possible. The SAP should include details of all of the analysis to be undertaken where possible. In particular, it is important to make a clear distinction between the aspects of a study that will be used for confirmatory hypothesis testing and the aspects that will provide data for exploratory analysis.

7. TIMING

The SAP and its related documents must be constructed and approved before:

- Unblinding any of the participants to be included in the statistical analysis, other than those unblinded for safety (when applicable)
- Database lock (on the same date as the database lock is acceptable)
- Any analysis is undertaken.

8. REFERENCES

SAP Template final analysis

9. APPENDICES

N.A.