

Training Statistical Programmers on SAP Review Skills

Sascha Ahrweiler, UCB BioSciences GmbH, Monheim am Rhein, Germany

ABSTRACT

A statistical analysis plan (SAP) describes the planned analysis for a clinical trial. SAPs must be carefully reviewed by statistical programmers working on the project for clarity and comprehension in order to construct analysis data sets and prepare planned Tables, Figures, and Listings. SAP review is a challenging task that requires both statistical expertise and abstract thinking skills, and is often times completed in the absence of available trial data. Statistical programmers must also develop critical evaluation and communications skills to prospectively identify issues and seek clarity with statistical colleagues on planned analyses on a trial or a development program. Statistical programmer education and training is highly variable and is derived from multiple disciplines, suggesting that comprehensive training on SAP review is an essential skill to ensure consistency and quality within the project. This paper provides training guidance on how a structured comprehensive SAP review and how awareness of statistics can be increased for the statistical programmer, with a resulting quality improvement of the statistical programming work products.

INTRODUCTION

Clinical studies are complex scientific experiments designed to provide evidence to answer questions regarding the safety and efficacy of products. Furthermore data generated as part of these clinical studies are used for regulatory applications and/or communications of study results in manuscripts, marketing materials, or other symposia. During the conduct of clinical studies the statistical programmer is presented with documents prepared to support the experimental design, data collection, and analysis and reporting. The statistical analysis plan (SAP) is one of these documents and it is one of critical importance. The SAP provides the statistical programmer with relevant information and detail on the scope of planned analyses, population definitions, and methodology on how prospective decisions are to be made for presenting study results.

Statistical programmers must develop document review skills that foster reading content, comprehensive understanding, and critical review of the SAP to support programming of analysis data sets along with analysis and presentation of study results. This paper provides insight on best practice and training methods for statistical programmers on how to review SAPs.

PURPOSE OF A STATISTICAL ANALYSIS PLAN

The SAP describes the planned statistical analysis of a clinical study as outlined in the protocol. In contrast to the protocol which outlines the analysis the SAP is a technical document which describes the statistical techniques for study analysis in detail. The SAP defines all the statistical output which will be included in the clinical study report. Shell tables, figures and sometimes listings are usually attached to the SAP although they should not be formally part of the SAP. The SAP and the annotated CRF are the documents which are most often used by statistical programmers to create their deliverables.

In general there are four different SAPs in the clinical development of a compound:

- **Statistical analysis plan for a clinical study** – describes the planned statistical analysis of a study
- **Interim statistical analysis plan** – describes the planned statistical analysis of an interim analysis for a study and therefore needs to address handling of partial unblinding issues in case of blinded studies. It also describes possible impact on the conduct and the complete final analysis like the possible adjustment of significance levels.
- **Data Monitoring Committee (DMC) statistical analysis plan** – modification of interim analysis used for DMCs and describes regular (e.g. monthly) data monitoring procedures for safety or efficacy questions. The DMC SAP also contains the DMC charter which clarifies exactly the names and responsibilities of the involved parties.
- **Integrated statistical analysis plan** – describes the planned analysis for an integrated analysis which is used for example in submissions. It defines the details of programming output for an ISS and ISE usually in one document.

PhUSE 2011

In the following the reader should keep the SAP for a clinical study in mind. The described methods can be easily adapted for the other documents as well.

Usually the SAP is written by the trial or project statistician by using a template. In most pharmaceutical companies the SAP will be written according to an available template which contains a standardized structure and which usually must not be changed at the highest section level.

In general the SAP should give more details about the planned statistical analysis than the protocol. At least the following content should be included in the SAP:

- Brief description of the study and purpose – e.g. description of conduct details which are important for analysis, study objectives and variables
- Statistical methods to be used – e.g. summary statistics of subject data (means, standard deviations, extreme values, counts with corresponding percents,), statistical tests (analysis of variance, t-tests,)
- Description of analysis populations – e.g. safety set, per protocol set, full analysis set,
- Data handling rules – e.g. imputation rules, algorithms for derived variables
- Complete table of contents with all TFLs to be produced by the statistical programmer as attachment to the SAP text (i.e. not as an appendix or any other formal and official part of the SAP)
- Shell TFLs to be produced by the statistical programmer to define the layout of the TFLs as attachment to the SAP text (i.e. not as an appendix or any other formal and official part of the SAP)

According to ICH E9 [1] the statistical analysis is planned a priori. Therefore the clinical trial statistician needs to ensure that the SAP is carefully reviewed and approved prior to unblinding of the study. For open label studies the SAP should be reviewed and approved prior to database lock approval.

The SAP is a document which is submitted to regulatory authorities as part of a submissions package. The SAP is also part of the appendix of a clinical study report. Therefore the SAP is critically important for documenting all the planned statistical analysis. The SAP is also stored in the trial master file and it is used during audits to check if statistical programming followed exactly the descriptions in the SAP. The SAP is meant to be a stand alone document. Besides the technical statistical details it should contain brief descriptions and summaries of the protocol. It should not only refer to the protocol.

IMPORTANCE OF AN EXTENSIVE REVIEW BY A STATISTICAL PROGRAMMER

The SAP is a technical document which describes techniques used in programs by statistical programmers. The SAP is the most important document for a statistical programmer since all deliverables from the statistical programmer are defined in this document. A critical review ensures comprehensive understanding of content relevant to statistical programming and will support improved quality of statistical programming work on a clinical study.

Therefore the awareness and correctness of the SAP is critically important for the statistical programmer in the following areas:

- Consistency with study protocol on study description and purpose
- Definition of analysis populations
- Data handling rules – imputation of missing data, algorithms for derived variables
- Appropriateness of described analysis for study
- Transferability of statistical methods into SAS® code
- Consistency between SAP text and TFL shells
- Use of standard TFL shells (if applicable)

HOW TO REVIEW A STATISTICAL ANALYSIS PLAN

The SAP is a document which is usually governed by an SOP and built from a standardized template. In most cases the SAP template is an associated document to the SOP, with a fixed table of contents (TOC) and mandatory sections with regard to the highest section level.

The statistical programmer must read, understand, and seek comprehension for all sections in the SAP. It is important to note that the subject matter expert of the SAP is the statistician. The statistician is fully responsible and accountable for this document. Statistical programmers should use their statistical background knowledge to review and give critical comments back to the author. Nevertheless the final decision about the implementation of the comments is at the discretion of the trial statistician.

As a statistical programmer you should note during the review any deviations from project standards or similar studies and ensure understanding of why and where these deviations occurred, as it will affect programming.

The SAP review should be performed in multiple stages. It depends on the expertise of the statistical programmer if these stages can be done at the same time. A less experienced programmer could do this review step by step. A more advanced programmer can do all stages within a single read of the document. Nevertheless every statistical programmer should keep the following points to check in mind:

- Correctness
- Consistency
 - a) ensure consistency to protocol
 - b) ensure consistency with regard to project standards
- Completeness – check if all TFLs mentioned in the SAP text are mentioned in the TOC and shells for the unique TFLs are available, check if all TFLs mentioned in the TOC and shell TFLs are described in the SAP text
- Degree of Details – are all necessary information mentioned and described in necessary details (e.g. baseline information, algorithms for derived variables, details in statistical models needed to setup SAS code,)
- Appropriateness – check if the planned statistical analysis is appropriate for the purpose (e.g. tables to describe the baseline characteristics of populations, tables for comprehensive description of the primary and secondary variables,)

In the next subsections examples for these points are explained in more detail.

CORRECTNESS

The correctness is easy to check and can be done at first reading of the document without any statistical knowledge. Therefore this check could also easily be performed by any other study team member and should not be in the focus of the statistical programmer.

Basically it needs to be checked if statements in the SAP are correct. This could vary from typographical or grammatical errors on the one side and to check for correctness of scientific descriptions on the other side. E.g. if a range is described as “100<DBP<80” this is incorrect and should be commented as “80<DBP<100”.

A check for correctness should also contain check for the correct use of SAP template. The document should be read for general errors or wrong content in single sections. Also due to the standardized structure the author should not be allowed to delete or add sections on the 1st level. The author should also not allowed to add a table of contents for all the TFLs in the SAP text. These items should be checked when checking for correctness.

CONSISTENCY

The checks for consistency are checks which are a little bit more advanced than the checks for correctness. Every statistical programmer should be able to perform the check for consistency. Basically these are checks to ensure consistency with the protocol on the one site and consistency with project specific standards on the other site. In other words the checks for consistency are actually comparisons of existing documents like protocol, project standard definitions or SAS programs with the SAP text.

In order to check the consistency with the protocol the following should be checked:

- Is the conduct described in the SAP the same as in the protocol?
- Are the objectives (primary and secondary) in the SAP the same as in the protocol?
- Are the variables (primary, secondary, supportive) in the SAP the same as in the protocol?
- Is the study purpose mentioned in the SAP the same as in the protocol?

It is important to note that inconsistencies between the protocol and the SAP are of course allowed. If for example the clinical study team becomes aware during a data review meeting that the planned analysis is not appropriate any longer, inconsistencies are allowed. In that case the decision to change the planned analysis should be carefully described in the SAP and should include the reasons for the change and impact on study result. In general the statistical programmer should note these inconsistencies and comment on them.

In order to check the consistency between the study SAP and the project the following should be checked:

- Did the author use available standard TFLs? If not is there a scientific justification for that change or was this just personal preference?
- In case derivation rules like relative days, drug relationship, etc. are mentioned, check if the mentioned algorithms in the study SAP are the same as used in the project. This could be either a comparison of two SAP documents, but could also include a comparison of an old SAS program with a new SAP text

It is important to note that inconsistencies are also allowed between the SAP and project standards. The statistical

programmer should be aware of these inconsistencies and give comments. If for example another study uses another definition of drug relationship classes the statistical programmer should comment on that. If this inconsistency was intentionally made it is important for the programmer to keep in mind that programs may need to be adapted.

Another important example is a deviation from available standard TFLs. If a statistician did not use available standard TFL shells the statistical programmer should comment. In general the statistician should give a justification for the deviation from a standard. When writing the SAP the statistician should not ask: "Can I use the standard TFLs here?", but rather ask: „Is there a justification not to follow the standard TFL". Using the standard TFLs is not a question of personal preference. If the standard can be used it has to be used.

COMPLETENESS

The check for completeness is a similar comparison as the check for consistency and is also of similar complexity. Every statistical programmer should be able to perform the check for completeness. Basically these are checks to ensure that the TFLs mentioned in the SAP text are completely covered by the TOC and the shell TFLs. In other words the checks for completeness are checks to ensure the within SAP consistency.

In order to check the completeness of the SAP the following checks are important:

- Are the TFLs mentioned in the SAP text explained in the TFL shells (and the TOC) in the same manner?
 - E.g. if the SAP text mentions something like "analysis will be done by visit and treatment group" do the shells present the layout also "by visit and treatment group"?
 - E.g. if the SAP text mentions that the age classes will be "<65 years" and "greater equal 65 years" does this correspond with the shells?
- Are the TFLs mentioned in the shell TFLs and the TOC also mentioned in the SAP text?
 - E.g. if the TFL shells contain AE tables for intensity, maximum intensity and relationship are all these described in words in the SAP text?
- Are the table of contents of the shell TFLs complete?

It is of major importance to compare the SAP text and the TFL shells. As already written above the TFL shells are not an official part of the SAP. The only document which will be in the appendix of the CSR and will be submitted to the authorities is the SAP text. This is the binding document for the analysis. Often statistical programmers just use the shell TFLs to setup their programs. In case of a mismatch between shell TFLs and SAP text it is always more important what is written in the SAP and not in the shells.

DEGREE OF TECHNICAL DETAILS

The check for the right degree of technical details is more complex than the previous three checks. This check requires more statistical programming expertise and some statistical knowledge. Basically these are the checks to ensure that the statistical programmer has enough technical details to perform the statistical analysis.

In order to check for the right degree of technical details it should be checked:

- Baseline variables
 - Are the baseline variables defined for all variables with a pre-dose measurement?
 - E.g. is it explained that the last value before dosing is taken as baseline value?
 - E.g. in a cross-over study is it defined that the last value before dosing in every period is taken as baseline value?
 - E.g. is a special visit identified as baseline visit and what happens in case some measurements are not done at this visit?
 - Are the definitions of baseline variables adequate and sufficient enough for statistical programming purposes?
 - E.g. if multiple measurements are used to define the baseline value is it described how to calculate the single value? This is often the case with ECG baseline values where the mean or median of three single measurements defines the baseline value.
- Data imputation
 - Do I understand the stated rules?
- Derived variables
 - Summary Scales
 - QTcF
 - Weighted means
- Summary statistics
 - Used population clear?
 - Denominator for percentages defined?
 - For change from baseline defined a-b vs. b-a?
- Statistical models
 - Enough details to set up the SAS code?

PhUSE 2011

- Will I find my own way from the text to the procedure?
- Structure of covariance matrix for repeated measurement
- Non-parametric vs. parametric approach
- Survival analysis: censoring mentioned?

Especially the last point “statistical models” needs some knowledge of statistical methods. Statistical programmers should not hesitate to ask questions for understanding to either clinical trial statistician or colleagues or managers with more advanced statistical expertise.

For the other bullet points the statistical programmer can benefit from many years of experience. The statistical programmer knows how difficult certain algorithms are to be programmed and should review the technical details in the SAP and check if it covers also special cases which could appear. If a statistical programmer does not have the necessary experience close collaboration with a more experienced programmer, mentor or line manager is expected.

Although the SAP is a technical document some statisticians prefer to have a special documents to explain technical details for statistical programmers. If that is the case and the statistical programmer agrees to that there are important things to consider:

- The technical details have to be available at the same time as the SAP
- The technical details should be added as an appendix to the SAP and need to be available in the electronic document management system at a later time

It is then important to review these technical details in line with this section.

APPROPRIATENESS

The check for appropriateness goes even further than the previous checks. In order to check the appropriateness of a SAP the reviewer needs some knowledge about statistical analysis and regulatory requirements.

Basically it needs to be checked if the described planned analysis is appropriate to answer the questions the study was planned for. The reviewer should ask if everything is described that needs to be described or if there is something missing in the planned analysis. This is clearly the expertise of the clinical trial statistician. Nevertheless a careful review of a statistical programmer ensures a better understanding of the statistical environment and it ensures a better final SAP.

For example it might be sufficient for a Phase I study to only prepare listings for medical history. Usually in Phase I studies healthy volunteers are included in the conduct. Therefore a table would not provide very reasonable information to describe the population characteristics. This is completely different in a patient study. In a patient study it is necessary to prepare a table for the medical history in order to describe the population appropriately due to the variety of pre-existing diseases. When reviewing the SAP text the statistical programmer should check:

- Do I have listings described where I need listings?
- Do I have tables described where I need tables?

Another example is the appropriateness of statistical models. Often people struggle between the difference of an “analysis of variance” (ANOVA) and “analysis of covariance” (ANCOVA). If in a study no influence of baseline information on the drug effect is expected an ANOVA should be used. If baseline information is considered to have an effect an ANCOVA should be used. For an ANCOVA it is necessary to name the covariable and to describe its handling. E.g. in an epilepsy study the number of seizures per day prior to first dosing might be an important variable for the later judgment of the drug effect. In this case the number of seizures at baseline and the change from baseline or the number of seizures during the conduct should be taken into the model. A statistical programmer should be able to understand the basic ideas of the study to understand the statistical analysis and to question the planned analysis. It is important to mention that a change of a statistical analysis will often result in a protocol amendment.

Regulatory requirements can be named as a third example for appropriateness of statistical models. According to the rules for the reporting of clinical study result it is important to give a certain set of results. The number of non-serious adverse events above 5% is one example which needs to be reported according to the FDAAA and which was usually not presented in the past. It needs to be ensured that all tables which are needed for result reporting are described in the SAP.

The check for appropriateness is also very important when it comes to the actual statistical analysis. There is a huge difference in the statistical analysis if a variable is dichotomous (e.g. responder yes/no), if it is categorical (e.g. response: none, mild, moderate, severe) or if it is continuous (e.g. response values can range between 0 and 100). For response variables a summary statistic would not be very informative. In this case a frequency table is more appropriate. In contrast a frequency table is unreasonable for a continuous variable. It should be carefully checked of which type the variable is and if the described analysis is appropriate for that case. The book Common Statistical Methods for Clinical Research [2] gives a good introduction in different statistical methods and the translation into

SAS code and might be very helpful for understanding the methods described in the SAP.

It is important to mention that checking for appropriateness needs many years of experience. It also needs statistical programmers who don't hesitate to discuss with the responsible clinical trial statistician. Of course the statistician will always make the final decision, but the statistician should be able to justify the decision in a discussion with the statistical programmer. The statistical programmer needs to keep in mind that in a complex field like the statistical analysis there are usually no dumb questions which statisticians dislike to answer. Discussions between statisticians and statistical programmers will help in better understanding the overall picture of a clinical study.

There are countless numbers of examples which could be given here to describe the check for appropriateness. In summary it is important to think outside the box and read between the lines. It should not be assumed that everything that is written in the text is sufficient. So always ask what is missing.

TRAINING OF SAP REVIEW SKILLS

The actual training of the SAP review skills should address all the different learning types of individuals. While some individuals prefer self learn trainings other individuals might prefer class room trainings. The following methods have shown to be successful:

- Guidance document on SAP review including checklists
- e-learning based on guidance document (including assessment)
- Initial class room trainings based on guidance document
- Mentoring by more advanced statistical programmers or line manager

A guidance document should be prepared to train statistical programmers in SAP review skills. This guidance document should explain the purpose of the SAP document. It all should give guidance about the different points (correctness, consistency, completeness, degree of technical details, appropriateness) as described earlier in this paper.

It is also helpful to provide checklists in the guidance document (e.g. "Are the described methods appropriate to analyze the primary variables?"). This checklist is meant to give the statistical programmer a structured way how he should perform his review. It is also helpful to keep in mind the different points to review in the document.

The guidance document should be available in an e-learning module with a corresponding assessment. Often e-learning tools are just used to track the training status of employees. It is important that the assessments are setup in a way which give the statistical programmer an idea about the best review practices. The main focus of the assessment should be that the trainee learns something and not that he fails the assessment.

For groups of new employees or for initial training sessions, class room training sessions should be held. These training sessions should be interactive and should contain a lot of examples which are discussed. The learning curve based on the review of good or bad written SAP examples and the corresponding discussion is priceless. At a first reading an SAP text snippet might appear correct, but if it is connected with programming issues it could show its weakness. This can be best addressed during class room training sessions.

Besides the above mentioned training types the most important way how to increase the review skills of statistical programmers is regular practice. Especially for the first reviews statistical programmers should be mentored by more advanced and experienced programmers. The trainee needs feedback if his SAP review addressed all important topics of the document. Often in the beginning of a career the statistical programmers do not really know what issues they will face in the later analysis. It is of critical importance that more advanced programmers mentor the less experienced programmers to avoid issues.

CONCLUSION

The review of a SAP is an essential part of the statistical programmers work. The detailed description in this paper gives guidance how statistical programmer should be trained to review the SAP. This ensures a consistent approach and common understanding. Based on all the detailed review practices it is obvious that a good review of all documents takes time. It is well spent time though since it helps to avoid issues when it comes to the actual programming process. A good document review will ensure a high quality product of all deliveries in the end. It also helps to clarify misunderstandings as early as possible.

REFERENCES

[1] ICH E9,
[HTTP://WWW.EMA.EUROPA.EU/DOCS/EN_GB/DOCUMENT_LIBRARY/SCIENTIFIC_GUIDELINE/2009/09/WC50002928.PDF](http://www.ema.europa.eu/docs/en_GB/document_library/scientific_guideline/2009/09/WC50002928.pdf), ACCESSED 26.AUGUST.2011

PhUSE 2011

[2] WALKER, GLENN A. 2002, *COMMON STATISTICAL METHODS FOR CLINICAL RESEARCH WITH SAS EXAMPLES*, SECOND EDITION, CARY, NC: SAS INSTITUTE INC.

CONTACT INFORMATION (HEADER 1)

Your comments and questions are valued and encouraged. Contact the author at:

Sascha Ahrweiler
UCB BioSciences GmbH
Alfred-Nobel Strasse 10
Germany – 40789 Monheim am Rhein
Work Phone: +49 2173 48 2438
Fax: +49 2173 48 1947
Email: Sascha.Ahrweiler@ucb.com
Web: www.ucb.com

Brand and product names are trademarks of their respective companies.