

PSI

Statisticians in the Pharmaceutical Industry

GUIDELINES FOR STANDARD OPERATING PROCEDURES for Good Statistical Practice in Clinical Research

Written by: PSI Professional Standards Working Party

Lesley Struthers, SmithKline Beecham Pharmaceuticals
Gary Atkinson, Pfizer Ltd
Les Huson, The Wellcome Foundation Ltd
Peter Morse, Scotia Pharmaceuticals Ltd
Anne Wiles, Brookwood Statistics Ltd

Edited by: Good Clinical Research Practices
Graham Ball, Wendy Bohaychuk

except Guidelines 12 and 13, which were written by:
Philip North, ASRU Ltd, and other members of the PSI Public Affairs Sub-Committee

All the Guidelines are regularly reviewed by the
PSI Public Affairs Sub-Committee

PSI

GUIDELINES FOR STANDARD OPERATING PROCEDURES for Good Statistical Practice in Clinical Research

CONTENTS

INTRODUCTION

- 1. CLINICAL DEVELOPMENT PLANS**
- 2. CLINICAL TRIAL PROTOCOLS**
- 3. STATISTICAL ANALYSIS PLANS**
- 4. DETERMINATION OF AVAILABILITY OF DATA
FOR ANALYSIS**
- 5. RANDOMISATION AND BLINDING PROCEDURES**
- 6. DATABASE**
- 7. INTERIM ANALYSIS PLANS**
- 8. STATISTICAL REPORTS**
- 9. ARCHIVING AND DOCUMENTATION**
- 10. INTEGRATED DATA SUMMARIES**
- 11. QUALITY ASSURANCE AND QUALITY CONTROL**
- 12. INTERACTION BETWEEN A SPONSOR COMPANY
AND A CONTRACT RESEARCH ORGANISATION
(CRO)**
- 13. FRAUD IN CLINICAL TRIALS**

INTRODUCTION

Professional codes of conduct for statisticians, such as those published by the Royal Statistical Society, stress general principles including the need for professional integrity, a primary concern for the public interest and the preservation of professional standards. However, there is a need for more detailed guidance in specific disciplines. Therefore, the PSI “Professional Standards Working Party” set out to provide guidance for the preparation of Standard Operating Procedures (SOPs) for Good Statistical Practice in clinical research.

The development of Guidelines for Standard Operating Procedures (GSOPs) appeared to be a natural step in the process of encouraging and publicising Good Statistical Practice in the analysis and reporting of clinical trial data. In preparing these GSOPs, the objectives of PSI are:

- to ensure that statisticians in the pharmaceutical industry are aware of the principles contained in published codes of conduct
- to provide guidance to encourage adherence to these principles in the application of statistics to clinical trials
- to publicise the principles of Good Statistical Practice for clinical research to other professionals, internal and external to the pharmaceutical industry
- to provide guidance in the preparation of SOPs to ensure compliance with the requirements of international Good Clinical Practice so as to satisfy regulatory requirements with respect to the collection, processing, analysis and reporting of clinical trial data.

The many varied roles statisticians fulfil within the pharmaceutical industry and related organisations (e.g. academic institutions and Contract Research Organisations (CROs)) preclude writing a definitive set of SOPs for all PSI members. Hence the decision was taken to develop GSOPs which would give detailed statements of the principles of Good Statistical Practice within the context of clinical research (thus building on the existing codes of conduct), but which would require that statisticians develop the guidelines into SOPs specific to their own organisations.

Conversion of these GSOPs into SOPs appropriate for a specific organisation requires expertise. Care should be taken that the level of responsibility given to statisticians is appropriate to the level and depth of statistical expertise within the organisation. The adaptation of the GSOPs into working SOPs should be carried out to reflect the structure and expertise of the organisation.

The GSOPs address the most common requirements of statisticians practising in the field of clinical research. The areas covered are not exhaustive. In some organisations it will be appropriate to combine some of the guidelines into one SOP or to divide other guidelines into several SOPs.

The manner in which procedures are implemented is as important as the procedures themselves if the procedures are to be effective within the structure of a particular organisation. SOPs which are properly implemented as part of an overall quality system not only improve the quality of the service or product but also improve the productivity and efficiency of the process. For example, good trial design and good case report form design can greatly improve the quality of the data collected in a clinical trial.

The PSI “Professional Standards Working Party” was set up in May 1990 and was responsible for developing 11 of the following GSOPs over a period of approximately 18 months. The development process included input from many members of PSI through written comments and workshop discussion sessions. The development process has continued since then with reviews and revisions of existing GSOPs, as well as the preparation of new GSOPs. This review process takes place over a two-yearly cycle, and is the responsibility of the PSI Public Affairs Sub-Committee.

The GSOPs and the Royal Statistical Society Code of Conduct are now augmented by other relevant documents such as the International Conference on Harmonisation (ICH) documents on Clinical Study Reports and Statistical Principles for Clinical Trials and the Association for Clinical Data Management (ACDM) Guidelines for Writing Standard Operating Procedures.

1. CLINICAL DEVELOPMENT PLANS

Objective

The objective of this GSOP is to define the responsibilities of statisticians with respect to the clinical development plan.

Every clinical trial should be undertaken for a scientific reason, as part of an integrated development plan and must be conducted ethically. It is recognised that in the case of industry trials there will generally also be a longer-term overall financial motive.

Procedures

- 1 Before the clinical phase begins the statistician should be aware of the pre-clinical analyses concerning the compound, the results of any other previous relevant trials and any safety or efficacy issues which may have an impact on the clinical development. Literature research should be performed to identify suitable efficacy variables and the magnitude of clinically relevant effects. The statistician should study the influence of prognostic factors and seek to improve understanding of the natural history of the disease/disorder/problem concerned, whenever it is possible and practical to do so.
- 2 Key decision points should be identified. Questions to be answered might include:
 - What is the optimal design of the trials?
 - When and how should efficacy be measured and are the methods reliable and sensitive?
 - What should be the duration of safety monitoring?
 - Which trials will be used to indicate activity of the compound and lead to scale-up of drug manufacture?
 - At what point will the optimal dose and frequency of the compound be identified?
 - How will information from the above be fed into the design of the Phase III trials?
 - What ongoing external trials are of interest?
 - When might those external trials impact upon the design of studies in the current programme?
- 3 The statistician should provide advice on the number of subjects required for efficacy evaluation in the clinical development of a product. This can be accomplished by advising on the appropriate set of trials to include in the development plan and estimating the subject numbers required for each trial. The statistician should also advise on the number of subjects required for safety evaluation, which will be performed on pooled data. The aim should be to optimise the development process and this will often mean exposing the minimum number of patients to trial therapy and conducting the minimum number of trials required to support registration for a product licence. Consideration should also be given to the inclusion of studies that enhance scientific understanding or the relative merit of a new therapy over an existing therapy, as well as addressing commercial issues.
- 4 Advice should be provided on the general design of clinical trials in particular therapeutic areas and on the primary variables that should be measured. In order to do this, the statistician should research the literature and understand the key efficacy and safety issues in the development plan.

The statistician should also give consideration to other design aspects such as secondary

variables, key safety variables, timing and frequency of assessment (particularly specification of the primary assessment time point), length of follow-up and planned database closure date.

- 5 The management of clinical data (see GSOP 6, **Database**) and data to be collected in all trials should be discussed. For ease of pooling data for any meta-analyses of efficacy and/or safety, data standards should be encouraged. Other staff (data management staff, clinical staff) also have responsibilities in this process. Common standards should be considered, such as dictionaries of medical terms, definition and timing of measurements, handling of protocol deviations.
- 6 Generic statistical analysis plans should be written for all clinical trials to outline the methods of statistical analysis that will be applied. Before data analysis, such plans should be developed into a comprehensive and detailed description of the methods to be employed (see GSOP 3, **Statistical Analysis Plans**).
- 7 Assistance should be provided for planning the timings and sequence of trials, the likely flow of data and the required resources. The dates and requirements for product licence applications, presentations, and publication strategy should be discussed.
- 8 The statistician should discuss with clinicians and investigators the results of analyses of trials as they become available, to evaluate how the results affect the clinical development plan. This plan should be updated as appropriate.
- 9 The statistician should be involved in the preparation of the dossier for submission to regulatory authorities and should be prepared to respond to questions from regulatory authorities.
- 10 The statistician should be involved in the preparation and maintenance of the Investigators' Brochure.

2. CLINICAL TRIAL PROTOCOLS AND CASE REPORT FORMS

Objective

The objective of this GSOP is to define the responsibilities of statisticians in the development, review and approval of clinical trial protocols and case report forms (taken to include all forms of permanent data recording including electronic source data).

The protocol should include a detailed and comprehensive statement of the objectives of the trial and the statistical methodology to be employed in order to meet these objectives.

Procedures

- 1 A statistician should review the entire protocol and should be authorised to comment freely on all aspects of the protocol, including amendments. The statistician should be aware of the relevant literature, the clinical development plan and the results of similar trials.
- 2 The statistician should contribute to the preparation of the protocol (and amendments) by writing certain sections or by providing information necessary for their completion.
- 3 The statistician should ensure that the protocol addresses the following items satisfactorily :
 - specific trial objectives with clearly defined and obtainable endpoints
If there is more than one objective or end point, they should be prioritised.
 - proposed design of the trial which is appropriate to the nature of the disease/disorder/problem and its indications, the definition and measurement of the endpoints, and practical constraints such as the availability of patients, the formulation of the trial medication and the availability of other resources
 - details of the study population with definitions of the inclusion and exclusion criteria
 - methods of recruitment and randomisation to treatment groups
 - sample sizes which are clearly specified and justified in terms of the trial objectives
The methods and/or computer package used for the determination of sample size should be referenced or documented, as should the estimates of any quantities used in the calculations.
 - nature and quality of the data to be collected
Details should be given of the methods and timings of data collection. The data collected must be relevant, measured in a consistent manner and appropriate in quantity.
 - methods of maintaining and breaking the blinded code and the reasons for the degree of blinding adopted
 - plans for formal or informal interim analyses or inspections of the data
Note that all unblindings of the data prior to finalisation of the trial represent an interim analysis.

Details of the composition and terms of reference of any independent data monitoring group should be specified. Rules for stopping the trial should be clearly specified (see GSOP 7, **Interim Analysis Plans**).

- plans for sample size reviews, with a view to the possibility of altering the planned number of patients, conducted in a blinded and non-comparative way
 - details of specific statistical hypotheses that are to be tested or specific parameters to be estimated in order to meet the objectives of the trial
 - Methods for accomplishing these should be stated.
 - multiplicity, in the sense of any or all of multiple trial objectives, measures or endpoints.
 - details of the rules that will be used to determine the evaluability of patients, especially in cases of protocol violations, withdrawals or dropouts (see GSOP 4, **Determination of Availability of Data for Analysis**)
 - the general strategy for dealing with each category of data
It is suggested that the data be classified under such headings as demographic characteristics, concomitant medication, medical history, dosing and treatment schedules, efficacy measurements, laboratory and other safety data, adverse event reports, etc. or by type of data (see GSOP 3, **Statistical Analysis Plans**).
 - compliance with the statistical aspects of the appropriate ICH guidelines
- 4 The statistician should ensure that the case report form is designed to collect data as required by the protocol. The case report form should only include relevant items that will be evaluated in the final report. Data management staff also has a responsibility to review the case report form.
- 5 The statistician should review and approve the protocol (and amendments) and case report forms. The approval should be documented. Data management staff also need to be involved in the review and approval process, at the very least as far as the case report form is concerned.

3. STATISTICAL ANALYSIS PLANS

Objective

The objective of this GSOP is to define the purpose and content of the statistical analysis plan.

The statistical analysis plan is intended to be a comprehensive and detailed description of the methods and presentation of data analyses proposed for a clinical trial, in order to avoid post hoc decisions that may affect the interpretation of the statistical analysis.

Procedures

- 1 The statistical methods to be used for the analysis of the trial data will be included in the protocol. The analysis plan will be finalised prior to data analysis and before treatment unblinding, to provide full details, even to the extent of including templates of tables, listings and figures to be presented in the statistical report. Any changes between the methods in the protocol and analysis plan will be explained in the analysis plan.
- 2 The statistical analysis plan should contain a statement of the objectives of the trial, as stated in the protocol. Also, it should refer to any relevant literature review, the clinical development plan and results from any other similar trials, as appropriate.
- 3 The statistical authorship of the analysis plan, version and date should be clear.
- 4 The statistical analysis plan should define the populations (e.g. intention to treat, as randomised, efficacy evaluable, etc.) to be used (see GSOP 4, **Determination of Availability of Data for Analysis**).
- 5 All primary and secondary end-point indicators should be clearly identified in the statistical analysis plan. If possible a single principal measure of efficacy should be identified.
- 6 The statistical analysis plan should specify the hypotheses to be tested and any parameters that are to be estimated in order to meet the trial objectives.
- 7 A full and detailed description of the methods of analysis and presentation should be provided for each type of data. Consideration should be given to the following:
 - methods for handling multiple observations
 - rules for calculation of derived variables including definitions that can be programmed from the data
 - use of baseline values
 - use of covariate data
 - analysis of subgroups

- methods for handling multi-centre data
 - treatment interactions, particularly with centre
 - interim or sequential analyses
 - rules for stopping the trial, and allowance for them in the analysis
 - methods for handling missing data
 - levels of statistical significance (one-tailed or two-tailed) and clinical relevance
 - methods for handling outliers
 - identification of fixed or random effects models
 - methods for handling withdrawals and protocol deviations
 - methods for point and interval estimation
 - approach to handling concomitant medications
 - methods for handling more than two treatment groups - multiple comparison methods
 - definition of the safety population
 - specification of computer systems and packages to be used for statistical analysis
- 8 Provision should be made within the statistical analysis plan for checking the statistical model and then for alternative methods to be used if the test assumptions are not met.
- 9 All tables, graphs etc. should be listed and/or templated, for both in-text and for the appendices. Rules for their format should be given, or other documents referred to. Any ambiguities should be clarified, such as the appropriate denominator for a percentage or the number of decimal places to be used, and notes for the programmer may be useful.
- 10 The analysis plan must be compliant with appropriate ICH guidelines, particularly E9 (Statistical Principles for Clinical Trials) and E3 (Structure and Content of Clinical Study Reports).
- 11 The analysis plan should be circulated for review and comment to clinical and medical writing departments, and any others who may usefully comment, such as data management personnel. The final version must be signed-off by the author, clinical and medical writing personnel.
- 12 The statistical analysis plan should be reviewed/updated immediately before the blinded code is broken (or before analysis begins in an unblinded trial).

- 12 Changes in the statistical analysis plan should be justified fully documented in the statistical report.

4. DETERMINATION OF AVAILABILITY OF DATA FOR ANALYSIS

Objective

The objective of this GSOP is to define a formal procedure for the selection of data to be included in statistical analyses (e.g. as randomised, efficacy evaluable, safety evaluable, etc.) prior to the unblinding of the clinical trial.

The aim of this procedure is to minimise the bias in the selection of the data to be excluded from analyses.

Procedures

- 1 Before a clinical trial begins, a document should be prepared which defines the rules to decide which data will be included/excluded from the agreed statistical analyses. This document should form part of the clinical trial protocol, either in the main body of the protocol or as an appendix.
Since the reasons for excluding patients from the analysis cannot be entirely foreseen at the time of writing the protocol, although general guidelines or well-known reasons can be described, evolving trial conduct may require some further definition of patients to be excluded, at the end of the trial, but prior to unblinding of the data. If anything, plans should err in the direction of caution and over-reporting, particularly from the point of view of safety.
- 2 Aspects to be considered (but not to be limited to) when determining the availability of data are as follows:
 - inclusion and exclusion criteria
 - acceptable timings for visit dates and measurements
 - compliance with treatment
 - the nature and quality of the data
 - incorrect randomisation/treatment
 - concomitant illness
 - concomitant therapies
 - withdrawal (treatment and trial)
 - patients entering the study more than once.
- 3 It is usual for there to be a number of patient populations used in the statistical analysis (e.g. safety, intent to treat, per-protocol). Each of these populations should be considered, as above.
- 4 The acceptable size (in number or percent) of the population should be documented e.g. full analysis and reporting of a per-protocol population less than 60% or more than 95% of the ITT

population is unlikely to be useful, (these percents are a guide only).

- 5 The document should state by title the person(s) responsible for determining the availability of subjects. The procedure should involve consultation between clinical, data management and statistical departments. Final responsibility of agreeing the patients to be included or excluded should lie with a clinician.
- 6 This document should be reviewed/updated after the collection of all study data, prior to unblinding of the data. Any changes should be justified in the statistical report.
- 7 The statistician should examine the consequences of the decisions taken regarding exclusions, in order to determine their effect on the robustness of the primary analysis, rather than to provide alternative analyses.
- 8 Reasons leading to unavailability of data should be documented in the data management report (see GSOP 6, **Database**).
- 9 Excluded data and the reasons for exclusion should be listed in the statistical report.

5. RANDOMISATION AND BLINDING PROCEDURES

Objective

The objective of this GSOP is to define procedures for randomisation and blinding in clinical trials and the documentation of these procedures.

The aim of these procedures is to avoid the introduction of systematic bias into the conduct of the trial. The methods apply to lists prepared before the start of the trial and dynamic randomisation methods such as Interactive Voice Response Systems.

Procedures

- 1 In the process of producing randomised code lists, documented procedures should exist to control access to the randomised code list. These procedures should include the following details:
 - method of production of the randomised code list
 - person(s) (by title) responsible for preparing and checking the randomised code list
 - distribution list of electronic and paper copies of the randomised code lists
 - storage and access to the lists
 - method of accounting for the randomised code lists at the end of the trial
 - method by which emergency access to the code for individual patients is to be organised during the trial.
(The ABPI guideline on emergency access to blinded codes requires that a company be able to decode the treatment within one hour.)
- 2 Consideration should be given to the use of stratified randomisation, for example to handle randomisation in a multi-centre study and to ensure balance for baseline prognostic factors.
- 3 Consideration should be given to factors that might be the subject of blocking. The clinical staff involved in the study should not be informed unless necessary for the conduct of the study.
- 4 The statistician should be aware of the procedures used in drug packaging and drug distribution to ensure that the randomisation codes are applied correctly.
- 5 Adequate steps should be taken to ensure that the treatments are indistinguishable in blinded trials.
- 6 The study protocol should define any individuals involved in the trial who should **not** be blinded to treatment. Individuals whose conduct could affect the interpretation of the results or the results themselves, should be blinded when it is practical to do so.
- 7 The circumstances for breaking the code must be clearly described in the study protocol.

- 8 When carrying out interim analyses of blinded trials, the integrity of the blinding of the trial should not be compromised. Only a person not directly involved in the running or conduct of the trial should have access to the randomisation code. Decisions on how the data will be analysed and presented should be made before the statistician conducting the interim analysis is unblinded (see GSOP 7, **Interim Analysis Plans**).
- 9 If the code is inadvertently unblinded during the conduct of the trial, this event must be fully documented in the statistical report.
- 10 All unblindings of the code for specific patients and for specific reasons should also be fully documented in the statistical report.
- 13 The randomisation schedule for each investigator, or other unit of randomisation, should be checked to determine that it has been followed. One method of doing this is to check whether code numbers have been allocated in chronological order. If the randomisation has not been followed the statistician should discuss any anomalies with clinical research associates to determine the possible source of the problem and should attempt to assess and document the effect of this on the analysis of the data. Unusual patterns of randomisation may indicate fraud and the statistician may need to undertake a statistical examination of the data for results indicative of fraud.

6. DATABASE

Objective

The objective of this GSOP is to define procedures which ensure the accuracy, validity and completeness of databases on which statistical analyses are performed.

Proper data management is required to ensure that a clinical trial database contains an accurate, valid and complete electronic record of the raw data, and that the database is secure. Much of the content of this GSOP is also covered in detail by the Data Handling Protocol Guidelines produced by the Association for Clinical Data Management (ACDM).

The following procedures might be the responsibility of either the statistician or the data manager responsible for the clinical trial. A number of the procedures are most likely to be the responsibility of that data manager.

Procedures

1 The statistician responsible for a clinical trial must ensure that the following are reviewed by appropriately trained personnel:

- the methods employed in creating the database
- the data or audit trail linking the final electronic database to the original case report forms.

The responsible data manager has a responsibility to ensure the integrity of the data.

2 The procedures to be followed in the preparation and documentation of the electronic database should be fully considered by appropriately qualified staff and should be described in a data management plan (also known as a data handling protocol). This should be written prior to receipt of the first case report form and updated as appropriate as case report forms are processed. The plan should be written in consultation with the project statistician and relevant clinical personnel. (See also point 6 of this GSOP.)

The data management plan should describe the following items:

- specification of the person who has approval authority for the database
- level of data monitoring performed prior to delivery of case report forms to data management
- details of coding systems to be used
- conventions for the naming of variables and files
- requirements for derived data fields (e.g. totals, averages, means, medians, etc.)
- specific details of automatic data validation/plausibility checks (i.e. checks of the data consistency using study-specific computer programs)
- handling of text in case report forms
- handling of missing data, giving due attention to the primary outcome and safety data
- procedure for query resolution, i.e. the generation and logging of queries
- documentation of all stages of data management
- details of quality control checks, taking account of which efficacy and safety data are critical
- details of the software and hardware being used
- details of any electronically transferred data

- interim data management status reports (for internal use, particularly in large, long term studies).
3. The data management plan, and description of any deviations from it, should be given to the statistician responsible for the clinical trial.
 4. For every database, there should be a detailed data trail allowing an external auditor to follow the progress of each data point from the original case report form, and any other original sources used, to the final electronic database. Items that should be considered in this process are:
 - internal logging and tracking of case report forms and other original source data
 - documentation of query handling procedures (internally and with the investigator)
 - techniques and documentation for editing errors on the database
 - processing and documentation of data errors after delivery of the database to the statistician
 - processing and documentation of data not included in the initial database design.
 5. Procedures should be defined to ensure that errors are not introduced into the database as a result of errors in computer programs. This includes the need to define programming standards. The following items should be considered:
 - (i) general system validation:
 - software to be used for data entry
 - structure and format of electronic data files
 - documentation of methods of software checking and action to be taken when errors are found.
 - (ii) study-specific checks:
 - checking of programs written for data entry (e.g. screen design)
 - checking of programs written to conduct data derivations and data validation/plausibility checks
 - checking of data manipulation programs written to transfer data from the entry format to the analysis format.
 6. Techniques should exist to ensure that the electronic database provides a complete and accurate representation of the raw data. (See also point 3 of this GSOP.) These techniques may consist of the following:
 - double data entry
Note that double data entry is the normal practice, but it may not necessarily be used. There should be an indication of how data comparisons will be made and who will make the decisions in the case of discrepancies.
 - programmed inclusion of electronic raw data
 - manual checking of the database performed after double (or single) data entry
Consideration should be given to the percentage and/or number of fields to be checked,

- the tolerable error rate, who will perform the manual checks and what action will be taken if the acceptable error rate is not achieved.
- manual or automatic coding of data in case report forms and the methods of checking the resulting data in the database
 - documentation of methods of database validation and procedures for correction of errors.
- 7 The integrity and physical security of the database must be protected. Techniques should be in place to prevent the creation of multiple versions of the database. Security procedures should consist of the following items:
- documentation of techniques for database 'locking' and security (including backups)
 - documentation and start time (e.g. before or after database locking) of an audit trail of database edits
 - documentation of controls on access to the database
 - archiving of the database (see GSOP 9, **Archiving and Documentation**).
- 8 Other issues to consider in preparing a trial database are:
- personnel (clinical, data management or statistical personnel) involved in making decisions with regard to identification of protocol violators
 - specification and preparation of data listings
 - timing of unblinding of the trial, if relevant.
- 9 The above assumes that there is a single database. However, the stored database may be passed over to another database for analysis, e.g. from ORACLE to SAS. In this case there must be full documentation and QC of the transfer process. Consideration should be given to which of the above apply to either, or both, databases. The statistician's input may be restricted to general issues and those issues related to the analysis database.
- 10 The data manager should ensure that a data management report is provided with the final electronic database. The report should describe the procedures and conventions that were followed in the preparation of the database, and include similar topics to those in the data management plan. The report should also include:
- a full field listing and description of the file structure of the electronic database
 - reference ranges and units for laboratory data
 - a list and brief description of all programs run on the data
 - level of errors found at each stage of checking the data
 - general comments on data quality and significant problems encountered with the data
 - a detailed list of any unresolved data queries
 - a statement of any queries/errors which have not been corrected on the database
 - a statement of the storage location of the electronic database.
- 11 The statistician should review the database prior to finalisation.
- 12 There must be a clear database closure process, including specification and proof of minimum data standards and of removal of write access and when the randomisation blind was broken. The process for re-opening and re-closure of the database must also be clear.

- 13 The statistician should inform the relevant data management staff of problems encountered with the database during analysis.
- 14 The statistician has a responsibility to document in the final statistical report any problems encountered with the management of the database.

7. INTERIM ANALYSIS PLANS

Objective

The objective of this GSOP is to describe procedures for the conduct of appropriate interim analyses that maintain the integrity of the trial and final statistical analysis.

An interim analysis is any examination of the data prior to locking the database of a clinical trial in which results (safety or efficacy or both) are evaluated by treatment group. The interim analysis plan should provide a comprehensive and detailed description of the methods of analysis and presentation of the data.

Unplanned interim analyses or analyses for management purposes only should be avoided. To cover the very rare situations in which such analyses may become necessary it is desirable to have a separate company SOP indicating the dangers and problems, and specifying the (high level) management approval procedures. All such analyses must be reported. Stringent analysis adjustments are necessary for all unplanned interim analyses.

Procedures

- 1 An interim analysis plan is a description of the proposed methods of analysis and presentation of the interim and final data. The interim analysis plan should be drafted concurrently with the protocol, or may be drafted after the protocol, provided it is complete before the start of the first interim analysis. It may be included within the protocol or may be a separate document, independent of the main analysis plan or part of it. Amendment of the plan should be justified and documented. Any unplanned interim analysis should be justified and fully documented in the interim and final statistical reports.
- 2 The interim analysis plan should clearly state the reasons (e.g. ethical, safety or to provide information to management) for the interim analysis.
- 3 The items to be considered and the content of the interim analysis plan should be similar to the statistical analysis plan (see GSOP 3, **Statistical Analysis Plans**). In particular, the interim analysis plan should address the following items:
 - variables to be analysed and data summaries to be reviewed (these should be as few as possible and consistent with the reasons for the interim analysis)
 - number of proposed analyses and when they will be performed
 - statistical model to be used for interim analyses
 - formal stopping rule or significance level adjustments
 - consequences for required sample size
 - the degree of unblinding and constraints on output to prevent the unblinding of specific patients (see points 4 and 5 of this GSOP)
 - process for dissemination of interim results (see point 8 of this GSOP)
 - method of handling additional data accruing after an interim analysis which has resulted in the termination of a study
 - consequences for estimation of all parameters at the final analysis.
- 4 The interim analysis plan should describe the degree to which the blind will be broken, how this

will be done and the steps taken to minimise access to unblinded information and treatment codes. It is preferable for the interim analysis to be carried out by a statistician not otherwise involved with the trial. Consideration should be given to whether the treatment groups should be partially blinded i.e. subjects allocated into correct treatment groups but treatments not identified.

- 5 For confirmatory trials using interim analyses to assess efficacy or safety, the use of an independent data monitoring committee (DMC, also known as a data safety monitoring board (DSMB)) is recommended. The responsibilities of the DMC should be agreed and documented. The interim analysis plan should be agreed by the DMC and Sponsor at the outset of the trial. Data supplied for each interim analysis, minutes of all meetings and data presentations should be archived.
- 6 Appropriate data management procedures should be defined.
- 7 Data reviewed at each interim analysis should be archived.
- 8 The dissemination of interim analysis results should be strictly controlled. Restrictions on circulation of results should be specified at the design stage. Dissemination should be restricted to the independent DMC, when one is appointed. In the study report it should be stated what information was disseminated, when and to whom.
- 9 The clinical trial report should clearly describe the design, conduct, results and consequences of all interim analyses (see GSOP 8), with discussion of the issues raised in this GSOP.

8. STATISTICAL REPORTS

Objective

The objective of this GSOP is to define the principal contents of a statistical report (hereafter taken to include also the statistical elements of an integrated clinical/statistical report), and procedures for the preparation, review and approval of a statistical report.

The statistical report is intended to provide a detailed description of the results of a clinical trial and the statistical interpretation of those results. It should also describe the statistical methodology employed in the analysis.

Procedures

- 1 The statistical report should be written by (or under the direction of) the statistician responsible for the trial.
- 2 The following details should be included in either the text or appendices of the statistical report:

Summary - a brief description of the study design, subject numbers, withdrawals, study populations, baseline comparability of groups, primary efficacy results, secondary efficacy results and main safety results. (This section is usually only included in integrated reports.)

Introduction - a brief description of the study.

Objectives – the study objectives, as stated in the protocol.

Study Design – a summary of the study design, including sample size estimation and method of randomisation.

Statistical Methodology - reference to whether all data collected are included in the report, description of and justification for any derived or transformed data used in the analysis, definition of the primary and secondary analyses, definition of the study populations, handling of missing or problem data, data used (e.g. observed, last observation carried forward), time point of analysis, evaluation of responses when the code was broken and methods of dealing with deviations from the protocol, documentation of discussions and decisions prior to code break (in an appendix), details and justification of statistical tests, null and alternative hypotheses, significance levels, tests of assumptions, estimation techniques, effects of any interim analyses, handling of multiple comparisons and centre effects, details of any deviations from the statistical analysis plan or planned analysis, specification of the versions of all software used.

Study Population - details of the number of patients recruited and randomised, the number of patients with data at each visit, reasons for all dropouts and withdrawals, the degree of compliance with treatment and the number of patients evaluable for analyses. Summary presentations of demographic data, clinical characteristics and initial severity of condition (for all populations e.g. ‘intent-to-treat’ (full analysis set) and ‘per-protocol’) should be reported.

Results - description for study populations of results from primary and secondary efficacy analyses and safety and tolerance analyses. The use of summary statistics (particularly confidence intervals), graphical data displays and the presentation of derived data should be considered. The order of presentation of results should be determined by the protocol objectives where this is consistent with any fixed format laid down by the company, e.g. for integrated reports. Details of the statistical tests should be reported. The details of statistical procedures should be described more fully in a statistical appendix if the analysis is complex.

Discussion and Conclusions - a brief discussion of the main statistical results with interpretation. Critical statistical features of the design should be assessed e.g. the appropriateness of a cross-over design. Factors such as the quality of data, problems with trial execution and external factors (e.g. weather in seasonal conditions) which may have influenced the trial should be considered. The sensitivity of the main conclusion to reasonable departures from underlying assumptions or inadequacies of the data should be discussed. Statistical features of the trial or its results, which are pertinent to higher level documents, should be highlighted.

Tables - labelled, referenced in the text and indexed. Units of measurement, explanation of rating scales, and the population on which the table is based should be stated.

Figures - labelled, referenced in the text and indexed.

Appendices - labelled, referenced in the text and indexed.

Data Listings -labelled, referenced in the text and indexed. Data listings should provide a clear and accurate representation of the data recorded in the study. Listings of derived data and patients included/excluded from analysis should be provided (at least for primary values, e.g. AUC (area under the curve), in the case of derived data). The randomisation list should be included.

- 3 Copies of the study protocol and/or the statistical analysis plan (and all relevant amendments of all documents) should be appended to the statistical report or, at least, held on file as should the data management plan or report and a protocol deviations document. Complex statistical methodology should be documented. The approach to the testing of the assumptions made in the statistical analysis should be documented.
- 4 The statistician should ensure that the statistical content of an integrated clinical/statistical report complies with the above guidelines, and that the clinical interpretation is consistent with the statistical findings.
- 5 The statistical report should be reviewed and approved by a second statistician (i.e. not the author) prior to issue. Appropriate internal quality control procedures should be followed (see GSOP 11, **Quality Assurance and Quality**).
- 6 The statistical report should be signed by the statistical staff responsible for the report (see point 5 of this GSOP).

- 7 The documents and audit trail indicating review and approval of the statistical report should be archived.
- 8 See also the following documents for further guidance:
- ICH - Efficacy Topic 3
Structure and Content of Clinical Study Reports
 - ICH – Efficacy Topic 9
Note for Guidance on Statistical Principles for Clinical Trials.

9. ARCHIVING AND DOCUMENTATION

Objective

The objective of this GSOP is to define procedures for producing archives containing sufficient documentation to permit the reconstruction of a statistical analysis and report from the original database. All aspects of clinical trials must be documented. The archived documentation provides the evidence of the conduct and findings of the project.

Procedures

- 1 A project log (or project work file) should be kept for all clinical trials and related projects, including small consultancy tasks. The log should contain a clear record of all activities that were performed during the project, if not specified elsewhere. Items that should be considered in the construction of a project log include the following:
 - the project log should be assembled in the early stages of the project (not at the end of the project)
 - the project log should be written in sufficient detail to allow a statistician to reconstruct the project
 - the project log should include:
 - general notes on all statistical aspects of the project including details of methods, conventions and procedures and correspondence on the conduct of the trial
 - a clear record of the database and details of how the database is labelled and stored (see GSOP 6, **Data Management**)
 - full details of all software used during the project.
- 2 All personnel working on a project should prepare documentation on the aspects of the project for which they are responsible.
- 3 The statistical component of the final study archive should include the following (if not stored elsewhere):
 - the project log (or project work file)
 - CVs and training records of statistics personnel
 - final protocol (with all amendments attached)
 - randomised code list and details of code breaks
 - statistical analysis plan
 - statistical report (or statistical elements of an integrated clinical/statistical report)
 - paper documentation of the location of electronic files
 - all relevant electronic data, programs, logs, output, emails and documents
 - documentation of QC procedures including sign-offs of report, programs etc.
- 4 Archives should be initiated at the start of a project and closed as soon as possible after the completion of a project and should be held securely. Material temporarily removed from the archives should be documented and tracked. If work is performed after the project has been archived, the additional documentation should be archived in the same format as the main project.

- 5 The archive may be almost entirely electronic. Paper copies of electronic documents do not need to be kept, but original paper documents that have been scanned may need to be archived.
- 6 The following documented procedures for archiving should be in operation:
 - a system of maintenance and storage of electronic media to allow for deterioration of the archives (e.g. floppy disks) and system hardware and software changes
 - indication of person(s) responsible for the archives and the access rights of others
 - length of time (since the last time the study was worked on) for which archives should be maintained.
- 7 Consideration should be given to retaining the archives throughout the lifetime of the product to which the report refers.
- 8 For data, which are, sent electronically from another source - e.g. laboratory data - an archive should be established at the original location of the data.
- 9 Before archiving, consideration should be given to material that needs to be removed for destruction.
10. The archive needs to be made secure against physical/environmental threats such as water damage/mould/fire/decay.

10. INTEGRATED DATA SUMMARIES

Objective

The objective of this GSOP is to define procedures to be followed when there is a requirement to combine and analyse data from a number of studies.

Data overviews may be required for submission of safety summaries to regulatory authorities or for preparation of overall safety, or efficacy, or other analyses involving data from more than one trial. These summaries use raw, subject level, data. The combination of summary data, used in meta analyses e.g. study treatment means, is not covered in this GSOP.

Procedures

- 1 If possible the overview should be foreseen and planned at an early stage of design (in the clinical plan) so that trials have common features which facilitate the overview e.g. common end-points, common terms, common times of assessment.
- 2 Quality control procedures need to be specified for checking data that are combined for analysis.
- 3 Analyses for data overviews should be carried out according to a well-defined 'overview analysis plan' which should be approved before any analyses commence. The plan should include the following items:
 - objective of the overview
 - rationale for the grouping of trials
 - sources of data
 - procedures for the selection of the data
 - choice of primary variables and subgroups
 - methods of analysis and reporting of the data.
- 4 The data overview may be a unique report or form part of a regulatory submission. The following items should be addressed in either document:
 - procedures used for selection of trials from which the data were sourced
 - discussion of any bias introduced in the selection of the trials
 - details of statistical methods employed
 - comparison of results from the analysis of the combined trials with the results of the analyses from individual trials
 - discussion of any biases arising from missing data or other inadequacies of individual trials
 - exploration of the sensitivity of conclusions to departures from assumptions (e.g. homogeneity of treatment effect) or decisions made (e.g. selection of trials or data).
- 5 All reports should be reviewed and approved by the statistician responsible for the data overview.

11. QUALITY ASSURANCE AND QUALITY CONTROL

Objective

The objective of this GSOP is to define how the quality of the statistical input to the clinical trial is maintained and controlled.

Quality assurance systems must be in place before clinical trials begin to assure the quality of the study. During and after the trial, quality control systems must be in place to control and verify the quality.

Procedures

- 1 To assure quality before clinical trials begin, the statistician should ensure that SOPs are relevant and up-to-date, that all statistical personnel are adequately trained, and that appropriate plans have been assembled for the development of products (see GSOP 1, **Clinical Development Plans**).
- 2 The clinical trial documentation should demonstrate that a statistician was involved in the design, review and approval of the protocol, the randomisation list, the case report form, the database design, the statistical report (or statistical elements of a joint clinical/statistical report), the final clinical report (or joint clinical/statistical report), and the clinical expert report.
- 3 All procedures employed in statistical analyses should be documented on hard-copy and/or in electronic format. The documentation should be sufficiently comprehensive to ensure that the analyses can be reproduced with the same results (see GSOP 9, **Archiving and Documentation**).
- 4 The clinical trial documentation should indicate that the statistician reviewed the presentations or interpretations of the data analyses (e.g. publications, study reports, integrated summaries, expert reports, promotional material, management meetings and discussions with regulatory authorities). The statistician should also have ensured that data were presented in an unbiased way, that any assumptions in the analyses were clearly stated and that the limitations of the methodology were taken into account.
- 5 The statistician should be responsible for ensuring the accuracy and validity of computer programs used in statistical analyses. The following methods of achieving this should be considered:
 - programs should be clearly documented with full use being made of commenting facilities within the program
 - formal flow charts should be prepared for complex programs
 - a record of the function of each program should be kept in the project log and in the program comments
 - a sample of the data derived by computer programs should be manually checked
 - all data sets derived from the study database should be checked to ensure that they contain the intended data
 - statistical analysis programs should be validated by analysing a dataset with independently pre-established results and validation documentation should be maintained

- programs written for the purpose of data tabulation, data listing or graphical presentations should be checked by comparing the output with the original database
 - codes used for data listings should be verified.
- 6 Validated generic code (such as pre-written and validated SAS macros) should be used whenever possible.
- 7 A full record of all computer programs should be kept in the project log (see GSOP 9, **Archiving and Documentation**). Commercial software should be fully referenced. The location of the programs and procedures should be clearly documented.
- 8 The statistical contribution to clinical trials should be independently audited. The audit report, the response to the audit report, and any action taken as a result of the audit should be documented in the project log (see GSOP 9, **Archiving and Documentation**).

12. INTERACTION BETWEEN A SPONSOR COMPANY AND A CONTRACT RESEARCH ORGANISATION (CRO)

Objective

The objective of this GSOP is to describe the management of the interaction between a sponsor company (the Company) - typically a pharmaceutical company - and a Contract Research Organisation (CRO) to which all or part of the statistical procedures in connection with a particular clinical trial have been contracted out.

Definition

In the broadest terms a CRO may be defined as an organisation which conducts all or some of the activities involved in the drug development process. A CRO is a business or an academically-based body (or both) whose primary objective is to undertake, on behalf of clients, scientific or medical trials, in total or in part, for payment on both a commercial and a contractor/contractee basis. The CRO may assume the regulatory obligations and responsibility for all or any part of a clinical investigation transferred to it in writing, including by electronic transmission. The transfer of obligation of a clinical study, or any part thereof, to a CRO, in no way relieves the Company of its responsibility for generation of quality data and analysis submitted to a regulatory authority.

Procedures

1 When selecting a CRO to undertake work on its behalf, in order to assess the suitability of a CRO for the intended work, the Company will need to consider various issues which may include, but may not necessarily be limited to:

- services offered by the CRO (e.g. CRF design, protocol development, study design, data entry, data management, statistical analysis, statistical reporting)
- security arrangements within the CRO
- ownership of the data generated by a study
- adherence by the CRO to Good Clinical Practice (GCP) standards
- existence of the CRO's own SOPs
- documentation by the CRO of its procedures
- qualifications and experience of the staff of the CRO, e.g. the CVs of any staff being supplied by the CRO to work on behalf of the Company, either at the CRO's premises or at the Company's premises
- expertise in the CRO in therapeutic areas relevant to the Company
- ease of access by the CRO to specialist expert help when required
- storage and archive facilities of the CRO
- computer software used by the CRO, including the version number, and the platform on which it is run
- quality standards of the CRO
- quality control and quality assurance procedures adopted by the CRO to maintain its quality standards
- the inspection status of the CRO with respect to the Company and the possible need for auditing of the CRO and the regularity of audit visits beyond the first one
- ability of the CRO to adhere to agreed time deadlines
- degree of flexibility offered by the CRO to be able to respond to changes in priorities and time schedules originating from the Company

- references on behalf of the CRO
 - the financial stability of the CRO
 - detailed costed proposal from the CRO for the particular, pre-specified piece of work to be undertaken, usually obtained as the result of a competitive tendering process.
- 2 The Company may wish to audit a CRO to which it has not previously contracted work, prior to inviting the CRO to tender for work on behalf of the Company.
- 3 In the tendering process for statistical reporting, the Company should provide the CRO with the following information as a minimum, in order to allow the CRO to prepare an informed proposal, except when the CRO itself is asked to develop and/or provide any of this information (other tenders will require different information):
- the protocol for the study (at least a draft version)
 - the CRF for the study (at least a draft version)
 - the number of subjects in the study
 - a detailed description of the statistical analysis and reporting to be carried out
 - any specific requirements concerning the format of the statistical report, including any relevant Company standard formats
 - timescale and deadlines.
- A meeting might be arranged before the tender is submitted between representatives of the Company and the CRO to discuss the Company's requirements. Whether or not there is such a meeting there should be a set-up meeting between representatives of the Company and the CRO which was successful in the tendering process, prior to the commencement of the CRO's work on behalf of the Company.
- 4 When supplying information to the CRO, the Company should require a confidentiality agreement to be signed by authorised people within both the Company and the CRO. The Company may wish to check that all the staff of the CRO are in any case bound by a confidentiality agreement as part of their conditions of employment.
- 5 In order to establish the relationship between the Company and the CRO for the particular piece of work to be undertaken, a formal legal contract should be drawn up for the work. This should cover the specification of the work to be undertaken with deliverables by each party, the timetable including milestones for both the Company and the CRO and the cost of the work and the payment arrangements. Work cannot start until the contract, or, at least, a letter of intent, has been signed.
- 6 The Company will need to know what indemnity cover there is for the CRO.
- 7 The Company should ensure that the work to be carried out by the CRO is specified in detail. Wherever there is a particular requirement by the Company (e.g. regarding Company format) this must be stated quite clearly to the CRO and agreed prior to the commencement of the work. If the CRO is to produce a statistical report on behalf of the Company, it will usually be helpful for the CRO or the Company to develop a detailed Statistical Analysis Plan before unblinding of the data, including templates for all the tables, figures and data listings to be included in the report, to be agreed by the Company and the CRO prior to unblinding of the

data.

- 8 The Company should assure itself that, in working according to its own SOPs, the CRO will also comply with the requirements of the Company's SOPs. If it is necessary in addition for the CRO to work according to any specific SOPs of the Company then these should be made available to the CRO, for use by any staff of the CRO who need to adhere to them.

- 9 There should be sufficient personal contact between the Company and the CRO during the course of the work to allow the work being undertaken by the CRO to proceed efficiently. This contact might include one or more project meetings involving relevant staff of both the Company and the CRO in addition to the set-up meeting. It should be made clear who are the appropriate contact people within the Company and the CRO, perhaps within speciality groups, in connection with a particular piece of work, and provision should be made for cover when any contact people are absent.

In ongoing contact with a CRO the Company will usually find it convenient to have the same contact person within the CRO (or persons if, for example, it is convenient to have a contact person for each separate function). Similarly, it will be helpful to the CRO to have the same contact person(s) in the Company over time.

- 10 In the case of a CRO supplying staff to work on-site at the Company's premises the Company may wish to interview the proposed member(s) of staff before the arrangement is agreed and finalised.

- 11 At the end of a piece of contracted work there should be feedback from the Company to the CRO indicating how well the work has been received and passing back any specific comments on the work.

- 12 The Company is likely to wish to establish ongoing performance monitoring of any CRO that it uses regularly. As part of this process regular audits of the CRO might be arranged.

- 13 Consideration should be given to the implications and responsibilities (of the Company and the CRO) of any subcontracting that might be proposed.

13. FRAUD IN CLINICAL TRIALS

Objective

The objective of this GSOP is to provide guidance on what actions the statistician should take when it is suspected that data from a clinical trial are fraudulent or is requested to investigate the data for fraudulence.

It is assumed that the company will have a clear policy on what actions to take if fraud is suspected including the statistician providing expertise. The role of the statistician is to provide supportive evidence of fraud and, possibly, to identify previously unsuspected fraudulent data.

Definition

Fraudulent data are data generated with the intention to deceive.

Procedures

- 1 If fraud is suspected the statistician must be informed and asked to review the data. Initially this should be done blind to the suspect centre.
- 2 If the investigator is proved to have supplied fraudulent data then a statistical review of the data may provide useful supportive information.
- 3 During statistical data review and analysis, good statistical practice necessitates that any anomalies in the data are investigated. If anomalies are found the statistician should discuss them with data managers, CRAs, and other appropriate personnel. These discussions may indicate that a data review for fraud should be undertaken, but even if there is no suspicion of fraud the statistician must still be satisfied that the data are explicable.
- 4 If fraud is suspected, the statistician should investigate the data with the intention of identifying characteristics that may support or, importantly, refute fraud. It is necessary that the subjects can be identified to the appropriate level of the fraud, in most cases identification by centre will be appropriate. As every case is unique, the data investigation must be broad and flexible. However, checks will largely focus on centre differences and may include: variability of the data, consistency, dispersion, digit preference, outliers, inliers, relationships between variables, relationships over time. Types of data reviewed, and checks used, are likely to include:
 - Randomisation: non-random order, poor balance
 - Dates: patient entry, visits (eg Sundays), differences between visits
 - Times: consistency
 - Single variables: Standard deviation, digit preference, outliers, change, SD of change
 - Multiple variables: relationships, residuals, leverage
 - Adverse Events: incidence
 - Treatment: different usage
 - Concomitant medications: study period differences.
- 5 A thorough data review needs the treatment blind to be broken. The impact of this on the analysis and the running of the trial must be considered and controlled.
- 6 The data review must be fully documented, with all output dated.

- 7 During the investigation of the suspected fraud the data will be in the study database. Once fraud is proved, fraudulent data must not be used in the reporting of the trial and may be removed from the trial database and kept in a separate database. Trial data associated with the perpetrator of the fraud, but not proved to be fraudulent, should be treated as fraudulent, unless requested otherwise by appropriate authorities.