MRC Clinical Trials Unit

Smarter studies
Global impact
Better health

UCL

STATISTICAL PRINCIPLES

VERSION 6.0

APPROVALS

Author	**Position**	**Signature**
Alex Szubert	Senior Research Associate	Alex Szubert Digitally signed by Alex Szubert Date: 2020.03.09 15:54:16 Z
Reviewer(s)	**Position**	**Signature**
Angela Crook	Senior Statistician	Angela Crook Digitally signed by Angela Crook Date: 2020.04.20 08:02:15 +01'00'
Ruth Goodall	Senior Statistician	Ruth L. Goodall Digitally signed by Ruth L. Goodall Date: 2020.04.20 13:40:33 +01'00'
Approver	**Position**	**Signature**
Fay Cafferty	Senior Statistician	Fay Cafferty Digitally signed by Fay Cafferty Date: 2020.04.20 15:07:23 +01'00'

All appropriate approvals must have been completed prior to uploading to SOPbox.

UPLOAD TO SOPBOX

Name	**Position**	**Signature**	**Date uploaded to SOPbox**
Cheryl Pugh	SOPbox Administrator	Cheryl Pugh Digitally signed by Cheryl Pugh Date: 2020.05.11 14:28:18 +01'00'	11-May-2020

The effective date of this SOP is the day on which it is uploaded to SOPbox and is available to use. This is the date associated with the signature of the SOPbox Administrator.

For the Revision History please see the Version History Summary in SOPbox.

STATISTICAL PRINCIPLES

TABLE OF CONTENTS

The following symbols may be used in this SOP:

Indicates a link to a related document

Indicates instructions to document trial-specific processes elsewhere

Throughout this document the MRC Clinical Trials Unit at UCL will either be referred to as ‘MRC CTU at UCL’ or ‘the unit’. In instances where neither read well in the sentence, ‘the CTU’ may be used.

1 PURPOSE

The purpose of this SOP is:

- To outline statistical principles and considerations in the design, conduct, analysis, and reporting of intervention studies (trials). Some but not all points are also relevant to observational studies.
- To define the role and responsibilities of the Trial Statistician and interactions with other members of the trial team.

Access to statistical expertise is essential throughout the design, conduct, and analysis of the trial.

Statistical considerations in the design and analysis of trials should broadly follow ICH Harmonised Tripartite Guideline: E9 Statistical Principles for Clinical Trials. All statisticians should familiarise themselves with this document.

ICH Harmonised Tripartite Guideline: E9 Statistical Principles for Clinical Trials
ich.org/page/efficacy-guidelines

2 RESPONSIBILITY AND ROLES

The following table lists the roles relevant to this SOP and a brief description of their responsibilities.

This SOP will be circulated for Read and Understood to all appropriate roles identified in the training matrix.

ROLE	RESPONSIBILITIES
Trial Statistician	• Each trial should have a designated Trial Statistician, with appropriate qualifications and experience, who assumes ultimate responsibility for all statistical aspects of the trial. • The Trial Statistician should be named in the trial protocol • The Trial Statistician will interact closely with other members of the trial team, including but not limited to the Chief Investigator, Project Leader, Trial and Data Managers and Data Programmers.
Delegated Statistician	• Some tasks may be delegated to other statisticians involved in the trial; the Trial Statistician should exercise appropriate judgement about which tasks should be delegated and should check that these tasks are performed appropriately and accurately. • The Delegated Statistician is typically responsible for maintaining the Statistical Master File, which should include a list of all statisticians who have worked on the trial giving details of names, dates, and their main roles and responsibilities. See Statistical Master File Contents List, MRC_CTU_TT_0131 • The Delegated Statistician is typically the signatory for the Trial Check List for Statisticians, MRC_CTU_TT_0132. • All statisticians are required to sign the Statistical Confidentiality Agreement, MRC_CTU_TT_0116, to be filed in their Training Folder.

Note that, in some trial teams, Trial Statistician and Delegated Statistician (respectively) may be termed Senior Statistician and Junior Statistician (respectively), or Senior Statistician and Trial Statistician (respectively). For the purpose of this SOP, the terms Trial Statistician (i.e. the statistician who assumes responsibility for all statistical aspects of the trial) and Delegated Statistician are used. For the remainder of this document, Trial Statistician can be replaced with Delegated Statistician as appropriate with roles as defined above.

3 PROCEDURES

Only broad general principles of statistical procedures are considered here. Additional details are provided in the associated working instructions and trial specific documentation.

Formal documents (e.g. statistical analysis plans) should be versioned in accordance with the Document Management and Version Control Policy.

3.1 COMPLEX TRIALS

For some trials (e.g. trials with shorter- and longer-term endpoints analysed at different times, or MAMS, platform or basket trials), it may be necessary for some activities to be per comparison or timepoint. If the trial is complex, the Delegated Statistician should discuss with the Trial Statistician whether or not there should be more than one Statistical Master File or more than one Trial Checklist for Statisticians.

Statistical Master File Content List

Trial Checklist for Statisticians

3.2 STATISTICAL INPUT TO TRIAL PROTOCOL

The following statistical considerations should be included in the trial protocol:

- A broad description of the method of randomisation. Overly detailed description should be avoided to prevent inference of the randomisation sequence.
- The outcome measures for the trial, distinguishing between efficacy and safety outcome measures, and primary, secondary and exploratory outcomes
- Sample size justification. In general, this should be based on the trial primary outcome measure
- The planned approximate timings of interim analyses, if appropriate
- The trial protocol should normally include broad stopping statistical guidelines rather than rigid statistical stopping rules (ie Pre-specfied stopping rules should be considered as guidelines), if applicable
- A summary of the methods of statistical analysis to be employed (to be expanded upon in the Statistical Analysis Plan)

As a minimum, the Trial Statistician should critically review the trial protocol throughout its development with special attention to:

- Details of the trial intervention(s)
- The trial design e.g. parallel groups, crossover, factorial
- The trial outcome measures. The number of secondary efficacy outcomes should be considered carefully with regard to subsequent analysis and reporting (including any adjustment for multiple testing). See NEJM statistical reporting guidelines (nejm.org/author-center/new-manuscripts)
- Patient eligibility/ineligibility criteria
- Methods for randomisation (where applicable)
- Methods for blinding (where applicable), including unblinding for a medical reason and the timing of unblinding at the end of the trial

- Measures to avoid bias and increase precision

Further detail is given in the MRC_CTU_TT_0196 Protocol Template and MRC_CTU_WI_0009 Random allocation methods: Theoretical Aspects. If applicable, refer to MRC_CTU_SOP_028 Blinded Trials.

3.3 SAMPLE SIZE CALCULATION

The sample size calculation must be independently replicated (or, where simulation was used, the code independently reviewed and executed) by a statistician who did not perform the initial calculation. This must be carried out before finalisation of the protocol, although it is recommended that this is carried out earlier in the development of the trial.

3.4 RANDOMISATION

Randomisation refers to the random assignment of participants to one of two or more groups which receive different interventions. Its main purposes are (a) to avoid allocation bias, (b) to minimise differences between the groups in terms of characteristics (measureable and unmeasurable) other than the interventions being compared that may influence clinical outcome (prognosis), and (c) to provide a basis for statistical inference.

Many different methods of randomisation have been developed, particularly to achieve balance across groups in terms of important prognostic baseline variables. This SOP does not mandate the use of any particular method, provided that the allocation to an intervention is unpredictable to the investigators who are entering patients into the trial and Trial and Data Managers and any other trial team members who are assessing patients' eligibility.

More information on the procedures for selecting and implementing random allocation methods is given in the Working Instruction on MRC_CTU_WI_0009. Random allocation methods: Theoretical Aspects.

For trials employing random permuted blocks, a randomisation list will be produced by the Trial Statistician or an independent statistician as appropriate for integration into the trial database. "Dynamic" methods of randomisation (e.g. minimisation) will, in other than exceptional cases, be achieved through the MRC CTU randomisation server, implemented by Data Management Systems.

Prior to recruitment of patients, the Trial Statistician (with Data Management Systems staff if appropriate) should, using dummy data, verify that the randomisation program is correctly allocating participants to the intervention groups. The Trial Statistician should assess other features of the randomisation program in conjunction with the Trial Manager. The amount of dummy data used should be proportionate to the complexity of the trial design (e.g. randomisation method, number of stratification factors/blocks, allocation ratio, number of intervention groups). As an additional safeguard, the Trial Statistician should check the randomisation procedure once recruitment has commenced.

Arrangements for manual randomisation (if necessary) should be established before the trial opens and will be detailed in a trial-specific working practice. A decision not to develop a manual randomisation system for a trial should be documented in the trial-specific documents.

Refer to MRC_CTU_SOP_78 Manual Randomisation – Principles and Procedures SOP.

In the case of a blinded trial, the Trial Statistician should liaise closely with the Trial Manager and Data Management Systems staff if appropriate concerning provision of the randomisation list to the appropriate organisations for packaging and labelling of the interventional product. Careful consideration should be given to procedures to limit the number of individuals who have access to information that would enable unblinding whilst ensuring patient safety can be maintained by swift unblinding if necessary. Arrangements for emergency unblinding should be established before the trial opens and will be detailed in a trial-specific working practice.

Refer to MRC_CTU_SOP_028 Blinding SOP.

3.5 STATISTICAL INPUT TO DATA COLLECTION AND HANDLING

The Trial Statistician should ensure that all data items required for verification of all eligibility criteria and analysis of outcome measures specified in the trial protocol are accurately captured on the Case Record Forms. Data items which are not strictly necessary for analysis or trial management should not be collected. It may be useful for a list of key data items for analysis to be produced (e.g. for data cleaning or specification of criteria for database lock).

See SOP MRC_CTU_SOP_0072.CRF Development and Maintenance.

The Trial Statistician should review the database requirement specification, in collaboration with the trial team, to ensure that the design of the trial's main database permits the efficient extraction of data in a format suitable for use in a statistical package (statistical analysis files).

Refer to the following SOP and working instructions:
MRC_CTU_SOP_010 Database Development
MRC_CTU_WI_0006 Obtaining data for statistical analysis from Macro
MRC_CTU_WI_0065 Obtaining data for statistical analysis from a CACTUS database
MRC_CTU_WI_0062 Dealing with Data not Collected by CRF
MRC_CTU_WI_0060 Data Querying Checklist for Statisticians

The Trial Statistician should develop batch programs which perform extensive range and consistency checks on the variables in the statistical analysis files. This includes verification of all eligibility criteria against corresponding data items, in addition to any eligibility checking which is incorporated within the trial database (e.g. verification of a blood pressure eligibility criterion against the actual systolic and diastolic blood pressure values). These checking programs should be run prior to each analysis (IDMC/interim or final), and ideally more frequently.

Further details are given in the Working Instruction: MRC_CTU_WI_0005 Good Programming Practice for Statisticians

Queries arising from these checks should be referred to the Trial/Data Manager for resolution. Please refer to the trial specific Data Management Plan. For an analysis for which a database lock is performed, the Trial Statistician will be responsible for defining when the data are clean and ready for database lock (in a database lock quality checklist). If a database lock is not performed and queries are resolved before an IDMC/interim analysis is completed then either (a) temporary corrections to the data (hard coding) can be written within separate statistical program(s) (with

suitable annotation) or (b) the statistical analysis files can be recreated from the updated trial database and the analysis re-run.

See the following Working Instruction and SOP:
MRC_CTU_WI_0018 Dealing with data queries identified by statisticians
MRC_CTU_SOP_018 Database Lock SOP

Errors in the data may remain despite efforts to resolve queries generated by range/consistency checks. Thus, sensitivity analyses should be considered in which (a) values outside clinically plausible ranges are set to missing, and (b) appropriate assumptions are made regarding inconsistencies between different data items. This is to avoid reporting illogical data and to reduce the influence of incorrect data on the results. These data alterations should be programmed within a single file and include appropriate explanatory annotation.

3.6 STATISTICAL ANALYSIS PLAN

A statistical analysis plan (SAP) should be signed off prior to data extraction for the analysis for the first IDMC/interim report which includes substantive data, that is sufficiently detailed to allow correct double programming of the primary outcome by an independent statistician. The outline Statistical Analysis Plan (SAP) given in the protocol is the first SAP version 0.1. It should be recognised that the SAP is a fluid document because (a) the analysis may depend on unpredictable aspects of the data (identified by blind review of the data) and (b) new analytical ideas may be developed during the course of the trial. If the SAP is amended, it should be up-versioned and signed off again prior to database lock for the final analysis (or any interim analysis which will be used for decision making).

The critical information contained in the outline analysis plan in the trial protocol (SAP v0.1) is not expected to change; any major changes (such as a change to the primary outcome measure) would need to be approved by the Trial Steering Committee and amended in the protocol.

The SAP should be prefaced with a brief summary of the trial, including its aims and objectives, the population being studied, the trial design, details of the sample size calculation, and method of randomisation.

Statistical Analysis Plan Template

All SAPs should be reviewed by <u>at least two</u> experienced statisticians (internal or external) which is likely to include the senior statistician if s/he was not the primary author. In selecting reviewers, consideration should be given to any specific methodological aspects which require review. Review should be documented in STOPOver, including the names of the reviewers and completed before database lock for final analysis.

Many trials will have substudies (eg translational studies, quality of life) which may be analysed after/separately to the primary outcomes measures. These may be detailed in separate SAPs.

3.7 STATISTICAL PROGRAMMING

Write permission to trial-specific statistical directories should be limited to the statisticians working directly on the trial. All files (including datasets and programs) associated with completed IDMC/interim reports and the final report should be set to read only as soon as the analysis has been completed.

Refer to the following SOP and Working Instruction:
MRC_CTU_SOP_018 Database Lock SOP
MRC_CTU_WI_0013 Storage and confidentiality working practice for statisticians

The computer software used for statistical analysis should normally be one of the major statistical packages e.g. Stata, SAS, R, as these are validated by the software developers before being released. If community-contributed packages (e.g. obtained from The Stata Journal or SSC) are used, the validation status of these should be considered, particularly if used for the primary analysis of the primary outcome. Publications should state the software used.

All programs should be structured, and contain detailed descriptions and comments throughout, to enable them to be easily followed and understood by another statistician. In particular, all programs should have header information and a brief description of what the program does. There should be clear documentation on statistical analysis file specification and/or the procedures for exporting data from the trial database.

Further details are given in the Working Instruction MRC_CTU_WI_0005. Good Programming Practice for Statisticians,

The primary analysis of the primary outcome as defined in the Statistical Analysis Plan (SAP, see **Section 3.6**) and other key outcome measures as appropriate, should be independently programmed from the cleaned analysis files by a statistician who did not perform the initial analysis. This should be carried out before completion of the first IDMC report/interim analysis which includes substantive data/is used for decision making, and on the final analysis file prior to the dissemination of any trial results. The verification of the consistency of the results from the two independent analyses will be the responsibility of the Trial Statistician. Particular care should be given to checking that **categorical variables have been correctly handled in the analysis**, particularly those defining treatment arm in trials with more than 2 arms, and that the **treatment label assignments are correct**.

3.8 INTERIM ANALYSES, INDEPENDENT DATA MONITORING COMMITTEES (IDMC) AND TRIAL STEERING COMMITTEES (TSCs)

IDMC/interim analyses are essential for monitoring the progress of a trial and for assessing data quality and completeness. In most trials, an IDMC will be established which will periodically consider data by arm from IDMC/interim analyses to assess the safety and/or efficacy of the trial interventions. The responsibilities and function of IDMCs are fully described in the IDMC Charter for each trial.

An outline of the analyses for the IDMC/interim report (e.g. list of analyses, shell tables) should be drafted by the Trial Statistician as early as possible and sent to the IDMC for their comment. IDMC/interim reports will typically consist of a closed (confidential) report, normally reviewed only by the independent committee and trial statistician(s), and an open report (which may be a sub-set of the closed report) which can be reviewed more widely by members of the TMT and TMG. For a blinded trial with both blinded and unblinded statisticians, the blinded statistician(s) will not review the closed report.

The IDMC/interim analyses will normally comprise a subset of those included in the SAP and should not be overly detailed. A guiding principle is that the IDMC/interim report should only contain information that would inform the IDMCs recommendation to the TSC.

The IDMC/interim reports (open and closed) should also briefly describe the design of the trial and contain information on the rate of accrual (against predicted accrual) and compliance with CRF return.

All data in the IDMC report should be presented in a 'digestible' manner in the form of summary tables and statistics and, wherever possible and appropriate, graphical summaries (input from the IDMC may be sought on what is most helpful). The interim data are displayed usually with minimal interpretation.

The IDMC may ask for further details or summaries of data as they regard appropriate. Requests for additional analyses from the IDMC not included in the outline IDMC/interim report (e.g. post hoc power calculations) should be critically examined by the Trial Statistician, documented in the meeting minutes (or report to TSC, if appropriate), and may be referred to the TSC if necessary.

The Trial Statistician, while respecting the independence of the IDMC, should draw attention to the dangers of over-interpretation of early data if this is relevant during IDMC meetings.

The closed IDMC/interim report should only be seen by IDMC members and the Trial Statistician(s) (and any observers for training purposes and/or associated administrative support staff). All these persons should undertake to respect the confidentiality of the report, and measures put in place to prevent unauthorised access to electronic and paper copies of the report.

See the Working Instruction MRC_CTU_WI_0013 Storage and confidentiality working practice for statisticians and MRC_CTU_TT_0278 DMC Template Charter, which includes confidentiality agreements for IDMC members and observers.

Analyses comparing outcome measures between randomisation arms should not be performed before the end of the trial except for IDMC reports. The final statistical report should note the dates of all interim analyses and the reasons for any unscheduled analyses.

3.9 DEVELOPMENT SAFETY UPDATE REPORTS (DSURs)

For IMP trials with a Clinical Trials Authorisation (CTA) within the European Union (MRC_CTU_SOP_003 – Regulatory Approval), annual submission of a Development Safety Update Report (DSUR) is required.

Development Safety Update Report Production details the responsibilities of the Trial Statistician

3.10 SUBSTUDIES AND RELEASE OF DATA BEFORE TRIAL CLOSURE

Proposals for sub-studies not already documented in the trial protocol should be approved by the Trial Steering Committee. IDMC approval should be additionally sought if the analysis of the sub-study is to be conducted before unblinding of trial participants.

The Trial Statistician should scrutinise any sub-study proposal and the request for release of any data before the trial is closed to ensure it compromises neither the main randomised comparison nor the blinding if applicable.

3.11 STATISTICAL REPORTING

Open and closed IDMC reports each comprise a single document detailing the results of all summaries/analyses performed. Final Statistical Reports (i.e. end of trial) comprise a single document detailing the results of all analyses performed at the end of the trial (eg an updated version of the last IDMC report) or a collection of such documents (eg if analyses of secondary outcome measures takes place at a later date and are summarised separately). In some circumstances, a collection of well-structured electronic files (eg Stata do and log files) documenting the analyses performed, together with the final publication(s), may be acceptable instead.

Trials which will be used for licensing applications will require a full Clinical Study Report as described in ICH E3 (in this case, production of the non-statistical sections of the Clinical Study Report may not necessarily be the responsibility of the Trial Statistician).

For all statistical reports (including IDMC/interim and final reports), tables should, whenever possible, be based on direct output from the statistical program. Ideally, report production should be automated (e.g. `putdocx` writes tables and images from Stata to a Word document, avoiding copy/paste error) or semi-automated (e.g. copy/paste of entire table as a whole without manipulating figures). The primary analysis of the primary outcome as defined in the Statistical Analysis Plan (SAP, see **Section 3.6**) and other key outcome measures as appropriate, should be carefully checked against the log files. When report production is neither automated nor semi-automated, all tables in the report should be carefully checked against the log files.

This must be documented, for example using the Statistical Report Checking Template.

Final statistical reports and IDMC/interim reports should be checked and endorsed by an appropriately experienced statistician other than the statistician who produced the report (a blinded trial with both blinded and unblinded statisticians may need a second unblinded statistician for this purpose), prior to their release. At a minimum, this will include reviewing the data for internal consistency and consistency with previous reports to identify anomalies. The primary analysis of the primary outcome as defined in the Statistical Analysis Plan (SAP, see **Section 3.6**) and other key outcome measures as appropriate, should be carefully checked against the log files. Particular care should be exercised for analyses, tables, and figures to appear in publications.

Any deviations from the statistical analysis plan should be described and justified in the final report of the trial.

Results of statistical analyses should be reported according to the EQUATOR Network guidelines that includes guidelines such as CONSORT. (See http://www.equator-network.org/) .

Presentations/publications should be checked against the final statistical report at the draft stage. Proofs of publications should be checked against the draft, or the final statistical report.

A summary of the final statistical report (which may comprise the trial publication or, where this is not yet available, a summary in the form of the “end of trial report” template) should be submitted to the main REC. For IMP trials with a Clinical Trials Authorisation (CTA) within the European Union (MRC_CTU_SOP_003 – Regulatory Approval), results must be uploaded to the European Clinical Trials Database (EudraCT). Uploading results is the responsibility of the Trial Statistician (ensuring results are uploaded on time is the responsibility of the Clinical Project Manager). For other

competent authorities, a summary of the final statistical report (as described above) should be submitted.

For details see the Trial Reporting and Communication SOP (MRC_CTU_SOP_014). If applicable, also refer to MRC_CTU_WI_0068 Uploading Trial Results to EudraCT.

The results of the analyses should be presented in a manner likely to facilitate the interpretation of their clinical importance. More emphasis should be placed on estimates of the magnitude of the treatment effects or differences and confidence intervals rather than significance tests. The phrase "P-values less than 0.05 (for example) are regarded as significant" should be avoided.

See MRC_CTU_SOP_014 Trial Reporting and Communication.