Department	Research & Development
Title of SOP	Statistical Analysis of Clinical Trial Data
SOP reference no:	R&D SOP 20
Authors:	G. Constable
Current version number and date:	Version 1, 03.11.2021
Next review date:	03.11.2024
Target audience:	Research, pharmacy and R&D staff
Distribution:	HUTH R&D internet Click on GCP SOPs for HUTH-sponsored CTIMPs https://www.hey.nhs.uk/research/researchers/gcp-sops-for-hey-sponsored-ctimps/
When this document is viewed as a paper copy, the reader is responsible for checking that it is the most recent version.	

Authorized by	Name	Sign	Date
R&D Director	Professor Thozhukat Sathyapalan		
R&D Manager	James Illingworth		

This page details the version history and the main changes made for each new version.

Version Log		
Version number and date	**Author**	**Details of significant changes**
Version 1, (03.11.2021)	G Constable	First SOP approved by R&D Committee on 03.11.2021.

Please note for definitions of acronyms refer to Appendix 1 of Management of SOPs. Refer to Appendix 2 of Management of SOPs for the standards to which clinical trials that investigate the safety and/or efficacy of a medicinal product are conducted.

All the HEY R&D GCP SOPs are available at:
https://www.hey.nhs.uk/research/researchers/gcp-sops-for-hey-sponsored-ctimps/

1 Introduction, purpose and who should use this SOP

- This SOP describes the process and control of statistical analysis of data obtained through clinical trials performed by or sponsored by Hull University Teaching Hospitals NHS Trust.
- It is crucial to the integrity of the trial that the clinical trial team have control over the statistical analysis of the interpretation and extrapolation of the data generated.
- Good clinical practise applies to the statistical analysis of trial data in that the processes followed are documented, pre-defined and are fully traceable/audit trailed.
- Every study protocol/SAP should have named study statistician who takes ultimate responsibility for the analysis of the data generated. The study statistician should have appropriate qualifications and experience.
- Statisticians should advise on formulating trial objectives, sample size, suitability of endpoint, randomisation and blinding.
- All trials should have a pre-specified statistical methodology documented for the trial either directly in the trial protocol or in a separate document such as the statistical analysis plant (SAP). This methodology should be comprehensive and provide a detailed description of the exact methodologies used, if any software is to be used and when analysis will take place (with regards to interim analysis).
- The plan should also include any planned handling of missing data, as is inevitable from missed patient visits, patients electing to discontinue the trial or patients being removed from the trial for a range of reasons.

Statistical analysis (including interim analysis) can be used to review or identify:

- Inconsistent data, data outliers, unexpected lack of variability and protocol deviations.
- Examine data trends such as the range, consistency and variability of data within and across sites.
- Evaluate for systematic or significant errors in data collection and reporting at a site or across sites; or potential data manipulation or data integrity problems.
- Analyse site characteristics and performance metrics.
- Select sites and/or processes for targeted on-site monitoring.

All processes used in statistics may be subject to inspection to verify the data has been accurately handled and reported in accordance with GCP principles which state that "*All clinical information shall be recorded, handled and stored in such a way that it can be accurately reported, interpreted and verified, while the confidentiality of records of the trial subjects remain respected" (Part 2(9) of Schedule 1 to SI 2004/1031).*

2 Scope

The SOP is applicable to all HUTH sponsored research studies and clinical trials. Where HUTH is co-sponsor on a clinical trials with another institution this SOP should be followed; any discrepancies between HUTH and the other co-sponsors SOP should be highlighted in the co-sponsor agreement.

This SOP is to be used as reference for the development and implementation of a study protocol with regards to its statistical validity and analysis. This includes the statistical aspects of the trial development, monitoring, interpretation and extrapolation of the trial data as well as any subsequent trial reports.

3 Statistical Input

- A suitably qualified and experienced statistician should be involved during the trial design process. This will provide justification for the number of subjects to be recruited for the trial.

- Trained statisticians should have extensive knowledge of trial designs and applicable statistical methodologies. They should be aware of new and evolving statistical methodology and designs used in clinical trials.

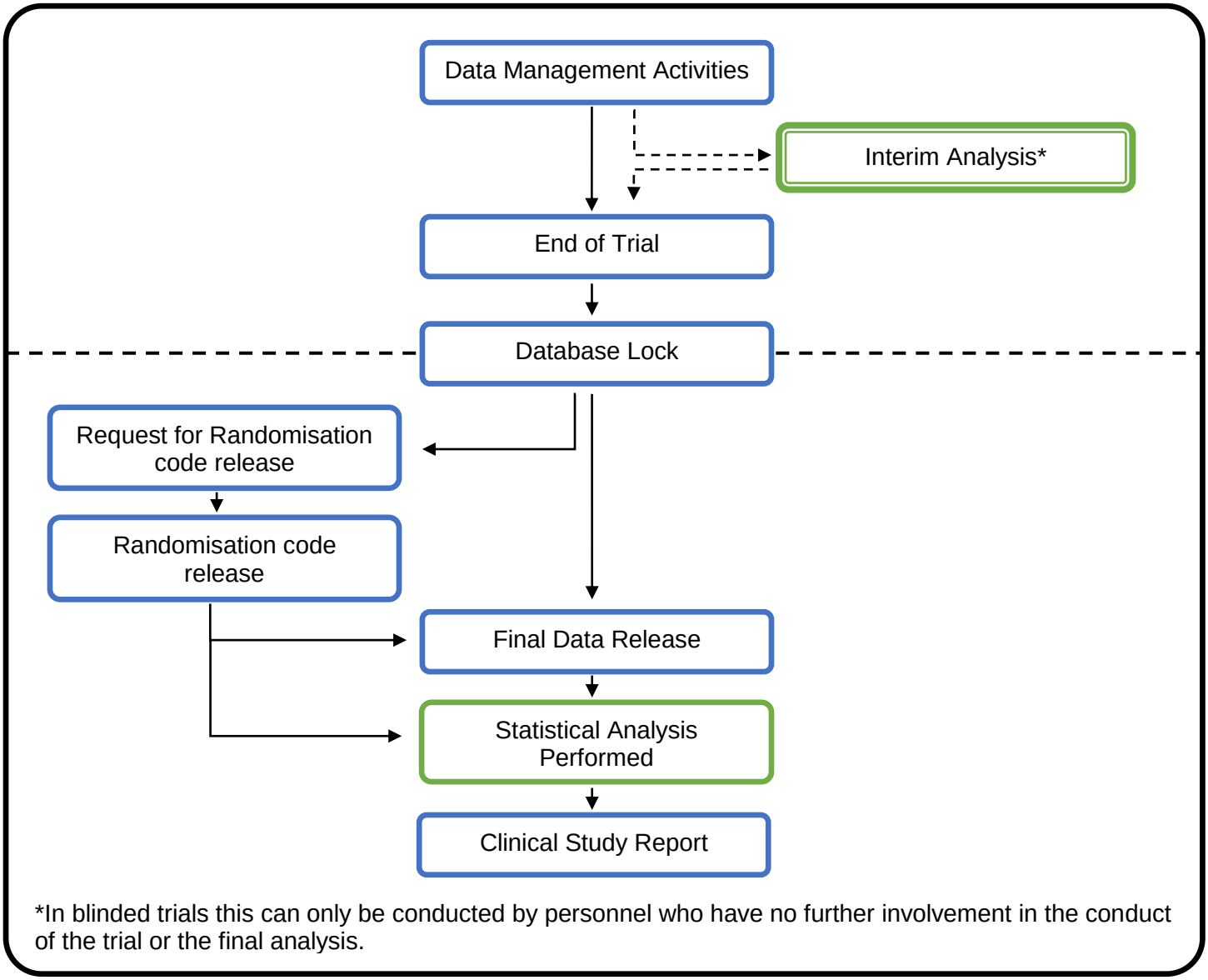
- This statistician must have current and relevant GCP training. This must be recorded and verifiable; ideally stored in a training record kept in the TMF.

- The statistician must have an awareness and/or training in the relevant guidance available; such as ICH Topic 9 –Statistical principles for clinical trials and the MHRA GCP Guide. It is always best practice to ensure current guidance is being used and is understood by all involved. Ideally this should be demonstrable in a user training record or similar.

- A review of all trial documentation (protocol, grant applications and manuscripts, for example) will require a statistician to review prior to submission to the MHRA, HRA and REC (or other relevant regulatory body).

- All statistical analysis should be performed by the named, experienced statistician.

A brief overview of the operations of a statistician during a clinical trial
(Shown in green)



*In blinded trials this can only be conducted by personnel who have no further involvement in the conduct of the trial or the final analysis.

4 Interim Analysis

The performing of interim analysis is necessary in some trial designs such as dose escalation trials. This must be detailed in the protocol as to what data will be used for the analysis, when it will be performed (in terms of triggered by a time-point or number of participants recruited) and with a full justification as to why the interim analysis is required.

A consideration required for the planning of interim analysis on a trial is that of unforeseen poor recruitment (particularly in regards to interim analysis arranged for a time-point rather than a recruitment target). If interim analysis is not to be performed at the time-point stated in the protocol due to, for example poor recruitment, the decision to cancel the interim analysis should be documented as a protocol amendment.
Including stopping rules into the protocol can assist in the generation of a protocol amendment as to why the interim analysis is not to be performed (for example, for futility, safety and efficacy).

Regarding interim analysis of blinded trials, the analysis must be performed by personnel who have no further involvement in the conduct of the trial or the final endpoint analysis. No unblinded personnel should be allowed to remain in a position to impact the conduct of the trial.

Unplanned or ad hoc analysis of unblinded trial data must not be performed as this risks the validity of the final report and has possible impact in the continuation of the trial.

5 Analysis

- Data entry must be completed by the investigator and follow the 'lock' procedure in line with R&D GCP SOP 12 End of Trial (with the exception of interim analysis which should have its own detailed procedure in the protocol/SAP regarding the locking of data used for analysis).

- All analysis performed should follow exactly as is stated in the protocol/SAP.

- All statistical analysis should provide an audit trail that is transparent as to what has occurred at each step and follows a logical flow which is repeatable by a secondary statistician, potentially in the absence of the original statistician.

- If further files are generated during the analysis process these should be saved as separate files and should clearly indicate to which stage of the analysis process they are. Ideally these files should be timestamped and username of the person saving it.

- The statistical analysis process should be clear and evident as to which dataset is used where and should be reproducible from the report alone.

- Quality control checks of the statistical analysis process is essential. These checks can be altered on a risk-based approach and should confirm data accuracy and adherence to the protocol. These should be detailed comprehensively in the protocol.

- The result of the statistical analysis could be directly written to the clinical study report (CSR) or produce a separate statistics report. The report should be version controlled and contain the relevant tables, graphs and data listings for use in the CSR.

6 Statistical Software

- If dedicated software is to be used, this should be clearly stated in the protocol and all audit trails available for the analysis performed.
- Any software used for statistical analysis of clinical trial data will require a prior risk assessment and validation.
- The use of dedicated software packages should always conform to data integrity principles and be assessed during review of the protocol.
- The level of validation required should be on risk-based approach; i.e. if the software is a bespoke, custom made software this will require a more thorough validation process than an 'off-the-shelf' software package from a reputable software manufacturer.
- Key features of the software should include unique user logons that are password protected, a comprehensive and easily accessed audit trail and safeguards in place against the corruption of data.
- Part of the validation of this software should include the preservation of the raw data during uploading from the study database (if carried out). Ideally this should be evidenced in a stress test of mock data being uploaded to ensure the data is not corrupted or otherwise altered.
- Reports generated from any analysis should be version controlled and protected.

NB: The level of software validation required for use in trials supporting marketing authorisations is much more thorough; assess the relevant regulatory approvals required in this case.

7 Suspicious Data

The use of a systematic, prioritized, risk-based approach to monitoring clinical trials is be employed on all trials to mitigate the risks (as per R&D GCP SOP 15 Monitoring) of suspicious data being generated which might be indicative of fraud. However with large sites collating large amounts of data, patterns might not emerge within the data until it analysed in its entirety.

Therefore statisticians and members of the data management team are in a position to identify possibilities of suspicious data generated by the trial. For example analysis of the data can highlight patterns or similarities between participants that could indicate the participant has been enrolled more than once.

Staff within these functions should be aware this aspect of the legislation and be trained in the relevant procedures for raising concerns they may have regarding the validity of the trial data.