UNIVERSITY OF WARWICK

STANDARD OPERATING PROCEDURE 08

Statistical Considerations

Version:	4.0
Approval Date:	29/05/2026
Effective Date:	12/06/2026
Review Date:	11/06/2028
Review Lead:	Ranjit Lall, Professor of Clinical Trials, Warwick Clinical Trials Unit (WCTU) WCTU Statistics Team
WCTU Reviewers:	Helen Parsons, Associate Professor, WCTU
Sponsor Reviewers:	Mathew Gane, Research Governance & QA Manager, Research & Impact Services (R&IS)
WCTU approval:	Natalie Strickland, Head of Operations, WCTU
Sponsor approval:	Carole Harris, Associate Director Research Culture, Governance and Compliance, R&IS

Revision Chronology:	**Effective date:**	**Reason for change:**
Version 4.0	12 June 2026	Biennial review: Minor text clarifications; addition of checklist C13, removal of QPulse
Version 3.0	23 February 2024	Biennial review: Minor text amendments and addition of data sharing considerations.
Version 2.0	20 October 2021	Biennial review: Change to new template structure. Expansion of background section and signposting of new SOPs
Version 1.6	24 July 2019	Biennial review: Inclusion of validation and programming checks. Removed overlap with SOP 21.
Version 1.5	11 July 2016	Biennial review: Addition in section 3.3.2. of requirement for blinded review of SAP; minor changes to text and expansion of range of software packages allowed in section 3.3.4.
Version 1.4	5 December 2013	Minor text changes to section 3.3.1 re protocol considerations and section 3.3.3 re documentation of analysis. Web links updated.
Version 1.3	21 May 2012	Biennial review: Web site link updated. Format change.
Version 1.2	1 April 2010	Biennial review: Web site link updated.
Version 1.1	31st January 2008	Biennial review: Format change.
Version 1.0	March 2006	

UNIVERSITY OF WARWICK

CONTENTS

1 Purpose and Scope

To detail the statistical input into the conduct of research studies including RCTs. It does not detail specific statistical methods, but general principles to be considered throughout a research study or clinical trial.

This SOP is applicable to statisticians involved in the development and analysis of research studies and to all University staff that are involved in designing or leading a research study including clinical trials.

For consistency throughout this SOP, we will refer to 'research study' to encompass all study types.

2 Acronyms

CI	Chief Investigator
CRF	Case Report Form
CONSORT	Consolidated Standards of Reporting Trials
DMC	Data Monitoring Committee
FRS	Functional Requirement Specification
ICH GCP	International Conference on Harmonisation Good Clinical Practice
RCT	Randomised Controlled Trial
R&IS	Research & Impact Services
SAP	Statistical Analysis Plan
SAMF	Statistical Analysis Master File
SOP	Standard Operating Procedure
TMF	Trial Master File
TMG	Trial Management Group
TSC	Trial Steering Committee
WCTU	Warwick Clinical Trials Unit

3 Definitions

SAP	Document detailing all planned statistical analyses to be conducted.
Trial Statistician	Person with appropriate training/experience responsible for oversight of, and/or undertaking the specialist statistical elements of a research study.
SAMF	A repository of all records related to the statistical oversight and conduct during a research study.
TMF	A repository of all essential records related to the oversight and conduct of a research study.
CRF	Forms on which individual participant data required by the research study protocol are recorded.
Oversight Committee(s)	A group of researchers, external to the trial team, whose role as a committee is to safeguard the interests of participants by independently assessing the integrity, safety and data produced by an ongoing trial. Commonly split into two committees: the DMC and TSC.
Randomisation	The process of assigning research study participants to groups using an element of chance to determine the assignments in order to reduce bias.

4 Background

Access to statistical expertise is essential before and during the entire research study process. ICH GCP Guidelines acknowledge that statistics is essential to both the design and analysis of clinical trials. For further advice for statistical considerations, the statistician should refer to the ICH Guideline E9 Statistical Principles of Clinical trials and E9(R1).https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf

5 Responsibilities

Trial Statistician(s)	• Ensure the research study is planned and conducted with due attention to all statistical considerations (see below) • Liaise with oversight committees to ensure data integrity and help monitor participant safety • Oversee and/or conduct the statistical analysis of the research study • Ensure appropriate reporting of statistical considerations and results • Ensure all essential records relating to the statistical considerations are filed in the TMF
CI	• Ensure statistical considerations are assessed appropriately • Liaise with the oversight committees and the statistician to ensure data integrity and participant safety • Review and approve the SAP
TM	• Work with the CI, statistician and programming team to develop the protocol, CRFs and database • Ensure all appropriate approvals are obtained, and essential records are filed in the TMF

6 When

It is recommended that you involve a statistician as early as possible when planning any research study, with statistical input provided throughout the entire study. As medical statistics is a specialised branch of statistics, it is further recommended that you consult an experienced or trained Trial Statistician when designing an RCT.

Statistical considerations must be assessed prior to an application for funding for elements including (but not limited to): the choice of the correct research study design; statistical properties of relevant outcomes; sample size and power. Statistical considerations should also be assessed throughout the research study, as detailed below. Elements range from the development of the protocol and case report forms, writing the statistical analysis plan, the conduct of analyses, write up of the results to the preparation of the data for archiving. The 'Statistics Road Map' (**G30**) provides a summary of activities requiring statistical input for different stages of a trial. It has been developed as a quick 'go-to- guide' for tasks and their associated SOPs. This document focuses on the key areas of statistical considerations.

7 How

For clarity, this SOP is worded such that a single statistician is responsible for all areas of statistical oversight. However, it is common practice that there are multiple statisticians working on a study as a team (e.g. Senior Trial Statistician and Trial Statistician). To enable clear delegation of duties, it is

recommended that statistical teams have a clear log of responsibilities throughout the study. Checklist **C13** (Statistical Processes for a Clinical Trial Checklist) can be used to define who is responsible for individual tasks, as well as recording when processes occurred and where appropriate evidence of the activity is stored.

7.1 Planning and protocol considerations

The statistician should be involved in the design of the research study and writing an application for funding as well as the protocol. Examples of relevant considerations that the statistician should have substantial input into includes:

- Review of existing evidence and justification for the research study
- Design e.g., parallel groups, cluster randomised, crossover design
- Criteria for evaluation of the objectives by the defined endpoints, e.g., response rate, quality of life assessment and methods of computation
- Methods for registration and randomisation of patients
- The population to be included in the research study and the rationale for the target differences between treatments which the research study is being designed to detect
- Measures to avoid bias, including blinding and completeness of data collection
- Sample size, taking into account clinical and scientific information and professional judgement on the clinical significance of differences that could be detected
- Methods for statistical analyses, including estimation of treatment effects and uncertainty, missing data etc.
- Interim analyses: whether any will be conducted and procedures for taking decisions based on them.
- Data sharing: consider if and how trial data will be shared. Will data be shared on request throughout the study, if a single data sharing pack can be created at the end of the study, etc.

The statistician must review the final version of the protocol and their review should be documented by signing the key document review/approval form and filed in the TMF. It is expected that any sample size calculations are checked by a second statistician prior to the protocol being finalised and that this can be evidenced in the reviewed documentation or via the approvals process used.

The statistician must also input into, review and approve the CRFs and the registration/randomisation procedures (as applicable). Review should be documented in the TMF.

7.2 Ethical approval and modifications

Research studies will require ethical approvals. Some committees will require a named statistician to be part of the research team, or a letter detailing that the research study has been reviewed and approved by an appropriately trained/experienced statistician.

Substantial modifications to the approved protocol and documents may be needed. To ensure that data integrity is maintained, and any impact of the modification is assessed, the statistician should be part of the review process for any modifications involving any statistical consideration. The modification process is explained in SOP 6: Modifications to Approved Study Documents.

7.3 Statistical Analysis Plan (SAP)

The SAP provides a comprehensive and detailed description of the methods that will be used to analyse the data and present the results of the research study. The SAP should be drafted following the guidelines detailed in SOP 21 'Statistical Analysis Plan' and, as appropriate, the associated SAP template document (**T21**).

7.4 Statistical Analysis Master File (SAMF)

All of the documentation, and programs for the analysis should be stored in the SAMF (electronic version). This will include the versions (where appropriate) of

i. validation checks,
ii. the SAP,
iii. randomisation (if generated by the statistician),
iv. testing of the randomisation,
v. Oversight committee reports and minutes which are confidential outside of the meeting (i.e. for 'closed' sessions of the DMC),
vi. analysis programmes/scripts for any purpose,
vii. locations of all datasets (or the datasets themselves) and
viii. interim analysis reports.

To enable reproducible reporting, programs/scripts should be written for the analysis wherever possible, and they should contain sufficient annotation to enable another statistician unfamiliar with the trial to reproduce the analysis. The statistician maintains the SAMF during the trial, and all documents should be stored in a location accessible only to the statistician and their appointed colleague that can access their encrypted files if the statistician is unavailable. After completion of the trial, the key statistical analysis documents (e.g., SAP, analysis programs/scripts and final analysis report) should be added to/signposted to the TMF for archiving.

A template structure for the Statistical Analysis Master File is available in the shared statisticians' folder on the CTU section of the M drive: M:\WMS\CTU\Statisticians\Stats Trial Master File Documents.

7.5 Statistical Analysis

The analysis of the research study must be carried out by an identified, appropriately qualified and/or experienced statistician and follow the SAP.

The statistician should be responsible for importing data into the validated statistical software package to be used for the analysis. SAS and STATA are the current standards for clinical trials, but other packages, such as R, SPSS, Excel, Minitab and Matlab may be used depending on the type of study and analysis required. The statistician should also ensure the latest version of the FRS from the programming team is used when analysing and interpretating the data. More details of the statistician's responsibilities can be found in SOP 14: 'Computer System Development and Validation'.

7.5.1 Data validation checks

Prior to any formal analysis, the statistician should obtain an initial data freeze of the research study database and place it on a secure specific directory. Where possible, the statistician should carry out validation checks on the data quality and integrity (e.g., range checks, outliers, missing observations). The statistician should refer any data queries arising during the analysis to the research study team for investigation, and resolution if necessary.

The study Risk Assessment and Monitoring Plan should include details of checks undertaken, how often they should occur and where documentation of the checks is held.

The research study team should be responsible for amending the current (dynamic) database and the statistician should, where possible after amendments, perform a re-freeze of the current database. Where time constraints are such that 'taking a snapshot' is not possible, for DMC for example, after discussion with the research study team, fully annotated hard coding of the correct data points can be

added into the statistical code as a temporary fix to ensure appropriate data representation within the frozen data, in addition to the correction being made in the dynamic database. More details can be found in SOP 15 part 4: Extraction of Data for Analysis and Data Lock.

7.5.2 Programming/scripts/code checks

All analysis programs/scripts must be reviewed by the statistician. This is especially important for user written code and code obtained from external sources. Where code has been written to create new variables derived from the raw datasets, the statistician must ensure that the derivation process has generated expected findings, e.g., values generated within expected ranges, for an expected number of cases.

On the first occasion of analysing the primary outcome, a second statistician should review the programming/code used to derive the primary outcome from the raw data. It may be appropriate to independently derive the primary outcome from the raw data if complex or extensive coding has been used to derive the primary outcome. If any discrepancies are identified, correction and re-analysis of the primary outcome will be undertaken by the statistician.

The review of the programming code should be documented on the Statistics Approval form (**T64**) and filed in the SAMF.

7.6 Interim and final analyses

Statistical results from clinical trials should be reported according to the CONSORT guidelines (http://www.consort-statement.org/) or the relevant extension appropriate for the trial design (see http://www.consort-statement.org/extensions).

The results of the analyses should be presented in a manner likely to facilitate correct interpretation of their clinical importance. Estimates of the magnitude of the treatment effects or differences and an appropriate measure of precision (e.g. confidence intervals) should be quoted.

Guidance on reporting other non-RCT studies can be found on the Equator network (Enhancing the QUAlity and Transparency Of health Research): http://www.equator-network.org/.

An interim analysis includes any examination of the data during the course of a research study for which results are presented for one or more treatment groups or collected data are used to decision making (e.g. stop for futility or re-estimate key parameters in an adaptive design).

The schedule of interim analyses to be conducted is usually determined by the Data Monitoring Committee; alternatively, it may be predetermined as part of the research study design.

The statistician is responsible for conducting interim analyses and producing open and closed (i.e., confidential) statistical reports as required. Output produced from statistical software may be manually copied into statistical reports or inserted automatically. Where possible the automated approach should be used. The open report should be checked by the TMG for typographical or analytical errors. The final version of the open report should be stored in the TMF. The final version of the closed report should be stored in the SAMF until the end of the research study when it should become part of the TMF.

Any changes to the originally specified schedule of analyses, e.g., interim analyses triggered by emerging data, must be justified and fully documented in the final statistical report.

The statistician should check that the statistical methodology stated within the trial SAP has been adhered to and report this on the statistical approval form.

The process of formally extracting datasets for interim and final analyses is explained in SOP 15 part 4: Extraction of Data for Analysis and Data Lock.

7.7 Data sharing

During and after the study, it is usually the responsibility of the CI and/or the study management team (e.g., the TMG) to review requests for data sharing. Requests must also be reviewed by the WCTU Data Sharing Committee. However, a statistician (or health economist or other suitably trained person) will need to be involved to ensure that data are shared in an appropriate format and a plan outlining how data is to be shared is included in the protocol. It may also be relevant to include in the SAMF (or elsewhere) a summary of the plans and management, or justification of why data sharing will not take place.

More details can be found in SOP 15 part 3: Sharing Data about how to facilitate data sharing, and a guide on performing anonymisation found in the guidance document (**G35**).

8 Associated Templates and Documents

T21	Template Statistical Analysis Plan
T09	Key Document Review/Approval Form
T64	Statistical Approval Form
G30	Statistics Road Map
G35	Data Sharing and anonymisation guide
C13	Statistical Processes for a Clinical Trial Checklist