UNIVERSITY OF LINCOLN

CLINICAL GOVERNANCE - STANDARD OPERATING PROCEDURE

DATABASE DESIGN

CG-QMS SOP CGD4

Version Final 1.0 Date 01 May 2021

Effective Date: 01 August 2021

Next review: 2 years

Author:	Dr Elise Rowan Clinical Data Manager
Approved by: Signature: Date:	

Version History	**Reason for change**

NOTE: All SOPs are subject to regular review.

Please ensure that the version of this SOP is the most up-to-date.

OUT OF DATE DOCUMENTS MUST NOT BE USED AND HARD COPIES SHOULD BE DESTROYED

1. PURPOSE

The purpose of this standard operating procedure (SOP) is to give general guidance relating to clinical databases/clinical database management systems and randomisation systems in relation to:

- System/database development
- System/database validation
- User training
- User support

2. SCOPE

This SOP is relevant to Chief Investigators, trial managers, data managers and statisticians who work on clinical studies and trials based within Lincoln Clinical Trials Unit (LinCTU).

3. BACKGROUND

3.1 Once the Case Record Forms (CRFs) for a clinical trial/study have been designed and approved in accordance with SOP CGXX Case Report Form, there is usually a need to develop a database or electronic clinical data management system in order to capture the data.

3.2 Systems can be standalone (offline) systems (or spreadsheets) for a single site study that is not a clinical trial or (more likely) online systems whereby multiple users at multiple sites can do data entry simultaneously. Commercially available online data capture systems are appropriate for clinical trials as these provide high quality audit trail tracking, allow for rigorous user access controls and are designed to meet well recognised national and international standards (see 3.6 below).

3.3 Systems need to be carefully developed and properly validated prior to use. Once implemented, users need to be trained and supported to use the database or clinical data management system successfully.

3.4 The main principles of this SOP should be followed when designing and managing any data management system whether it be a simple spreadsheet, complex online database or randomisation system. During the design stage it is essential to consider how to securely store and manage access rights to patient identifiable information; guidance is given for this in SOP CGD2 Data Management and CGD1 Data Protection and Confidentiality.

3.5 This SOP does not go into specific detail relating to specific software packages.

3.6 Databases for clinical trials should be produced in accordance with UK and EU standards. International standards may be required for multi-centre international studies. This should be considered when choosing commercially available Electronic Data Capture systems.

UK STANDARDS

Data Protection Act 2018 -this is the UK implementation of the General Data Protection Regulation (GDPR).

UK Policy for Health and Social Care Research - This policy framework sets out principles of good practice in the management and conduct of health and social care research in the UK.

EUROPEAN (EU) STANDARDS

GDPR stands for General Data Protection Regulation. It is the core of Europe's digital privacy legislation.

ICH GCP is an international quality standard for Good Clinical Practice (GCP) that is provided by the International Conference on Harmonisation (ICH)

INTERNATIONAL STANDARDS

ISO 27001 is the international standard that describes the specifications for implementing an information security management system (ISMS).

<u>USA Standards:</u>

- 21 CFR Part 11: is the United States Food and Drug Administration's (FDA's) regulations for electronic documentation and electronic signatures. It outlines the administration of electronic records in a medical device company's quality management system.
- FISMA compliance is data security guidance set by The US Federal Information Security Management Act (FISMA) and the National Institute of Standards and Technology (NIST). NIST is responsible for maintaining and updating the compliance documents as directed by FISMA.
- The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a US federal law that required the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge.

4. CROSS REFERENCES

4.1 CG-QMS SOP CG04 Protocol Development

4.2 CG-QMS SOP CG08 Trial Initiation.

4.3 CG-QMS SOP CGD1Data Protection and Confidentiality

4.4 CG-QMS SOP CGD2 Data Management

4.5 CG-QMS SOP CGD5 Case Record Form

4.6 CG-QMS SOP CGD6 Randomisation

4.7 Trial / Study Protocol

5. PROCEDURE

Prior to development the following should be in place:

5.1 Study/Trial Protocol; the trial protocol or related Data Management Plan should outline how data for the project should be managed and who should be able to collect, process, access and analyse it. There should be a clear plan or diagram within the protocol to indicate how participants progress through the study/trial and the number of visits and type of assessments that should be completed with them at each stage.

5.2 CRFs, standardised questionnaires and test instruments (with scoring algorithms) and lists of any source data to be captured should be described in the protocol and have been finalised in accordance with SOP CGD5 Case Record Form should be available in paper format prior to developing the database systems for the project.

5.3 Such documents should serve as the basis for the "data dictionary" for the study/trial and should be shared with the data manager who is tasked with developing the database.

5.4 The protocol should detail the exact process of randomisation in accordance with SOP CGD6 Randomisation.

5.5 There should be a paper record detailing the inclusion and exclusion criteria and the exact information needed to randomise a patient into the trial/study – typically this will be contained in the Trial Protocol.

5.6 The Trial Statistician (where appropriate) or CI should arrange to work with the data manager to ensure that the randomisation system is developed according to the randomisation method outlined in the protocol.

5.7 Consideration should be given to the type of database software to be used; for all clinical studies sponsored by University of Lincoln, this will be a University of Lincoln ICT Department approved Electronic Data Capture (EDC) software or computing package. For clinical trials or studies supported by LincTU, projects should be developed (prior to funding) with a view to using the pre-approved Electronic Data Capture (EDC) software available via the Lincoln Clinical Trials Unit (LinCTU).

5.8 It is expected that database designers will select software appropriately, considering what is recommended by the ICT Services at University of Lincoln and using, any relevant security and audit

trail features of these packages as required by the Sponsor or Funder, in order to meet any trial/study specific requirements* relating to data handling, storage, security and processing.

*Study specific requirements for database software could include but may not be limited to, the following UK/EU regulatory compliance standards: The UK Policy for Health and Social Care Research, ICH Good Clinical Practice, EU Clinical Trial Directive and Medicines for Human Use (Clinical Trials) Regulations and subsequent amendments, GDPR, Data Protection Act. Certain software packages may also allow for compliance with other international and USA data processing standards such as ISO27001, 21 CFR Part 11, FISMA, HIPAA (see background section).

5.9 During the development stage, the database developer/manager should work with the Chief Investigator (CI), study statistician and study team (as appropriate) to ensure that all information required for the database or randomisation system is captured appropriately. This can be achieved by regularly reviewing development progress and demonstrating the randomisation system or each electronic CRF as it is completed. A full audit trail should be retained.

DATABASE / RANDOMISATION SYSTEM VALIDATION

5.10 Electronic data capture systems shall reflect the layout, design and content of data capture documents/CRFs/proformas.

5.11 Electronic data capture systems shall allow data validation, range checks, consistency checks and developers should thus employ the use of these types of features to ensure accurate and high-quality data entry:

- Defined field types
- Requirements for unique identifiers
- Mandatory input
- Input masks (format limits)
- Range limits
- Drop down menus, list boxes, combo boxes, radio buttons
- Referential integrity

5.12 Where data is captured at more than one time point for each participant, the system shall have the flexibility to capture the name description of the scheduled study time-point (e.g. baseline, visit 1, etc) as well as the actual date and if necessary, time carried out.

5.13 If necessary, data entry checks or double data entry will be considered for critical data as defined by the protocol and staffing capacity/availability for the study.

5.14 Prior to release, it is recommended that dummy data is entered into the system to check that all the features of the electronic data entry forms work correctly.

This is essential for drop down menus and selection boxes.

It is also important to check that underlying data is stored correctly so that it can be used for later analysis, (or example, with the correct format and number of decimal places).

DATABASE/RANDOMISATION SYSTEM RELEASE

5.15 Once the database/randomisation system is finalised, a formal file note should be created in the Trial Master File (TMF) which records:

- The database/randomisation system name, version number and date finalised
- The software used to create the database/randomisation system
- Reference to the paper CRFs/source data and/or data dictionary used to create the database/randomisation system (these may be cross referenced in other sections of the TMF)

- Confirmation that the database/randomisation system has been approved by the CI (this can be a formal signed form or a copy of an email approving the system)
- The name and contact details of the data manager responsible for looking after the database/randomisation system and managing user access rights
- Details of which members of the research team have access to the database/randomisation system (at each site as required)
- Where the system is located (e.g. University Server, Cloud based, online web address)
- Any changes to any of the above as the project progresses.

USER TRAINING

5.16 The CI should work with the Data Manager to establish a list of all staff members who require training to use the database or randomisation system. Trial Training should be in accordance with SOP CG08 Trial Initiation.

5.17 Access rights should be assigned to members as appropriate depending on their role in the study/trial and need to either randomise, access or enter certain types of data as detailed on RF CG08-RF01 Study Delegation Log .

5.18 A study specific training plan should be developed which could include but not be limited to, the following:

- A slide show which demonstrates key features of the system.
- Written step by step instructions for logging on and accessing the system.
- Face to face training.
- Regular site training and monitoring visits by the data manager or trial manager.

5.19 A formal record of who has received the database training, by whom it was given and when should be kept in the trial master file along with copies of any instructions/training slides in accordance with SOP CG07 Trial Initiation.

USER SUPPORT

5.20 Each member of the study team should be given:

- Work contact details (email and telephone) of the data manager in case they need further database or randomisation system support once the project is underway.
- A copy of any written instructions or training slides relating to database/randomisation system use.

5.21 The Data Manager should be responsive to requests from the team relating to any clarification / explanation of database/randomisation system use. Such requests should be dealt with as promptly as possible during working hours so as not to cause significant delays to the project.

6. FLOW CHART

None required.