Research Governance
and Integrity Team

Database Lock

SOP Reference: RGIT_SOP_046	
Version Number: 3.0	
Effective Date: 26 Nov 2024	Review by: 26 Nov 2027
Author: Rinat Ezra, Clinical Trial Manager	
Approved by: Keith Boland, Senior Clinical Trial Manager	Date:

Version	Date	Reason for Change
Version 1.0	29 Jan 2020	1st Edition
Version 2.0	19 Oct 2020	Scheduled Review JRCO name change to RGIT.
Version 3.0	26 Nov 2024	3yr SOP review

IMPERIAL

TABLE OF CONTENTS

IMPERIAL

Research Governance
and Integrity Team

NHS
Imperial College Healthcare
NHS Trust

1. PURPOSE

The purpose of this Standard Operating Procedure (SOP) is to describe the procedure for personnel involved in the process of locking a clinical trial database, including CIs, Statisticians, Programmers, Trial/Data Managers and Coordinators. It also covers the procedures required should a hard-locked database need to be unlocked.

For CTIMP Trials it is mandated for the trial personnel to use the authorised in-house system (i.e. previously InForm now transferred to OpenClina for the College Sponsored CTIMPs) for data capture.

For non-CTIMP studies the College provides the REDCap system to use but it is not mandated to do so.

This SOP is primarily to provide higher level/oversight information for Trial database locking.

2. INTRODUCTION

The process of locking a clinical trial database is an action taken to prevent further changes to the database. A database is locked after review, query solution to the best of the resolution possible/available and confirmation that data captured on the database is ready for analysis.

In order to prepare for an interim or final analysis of a clinical trial, the data on the dataset(s) to be analysed must be finalised. Once the entered data on a database is complete and clean the database is locked to prevent further changes to the data and the analysis may be performed.

3. DEFINITIONS AND ABBREVIATIONS

The headings below contain the definitions of terms of meaning of abbreviations used within the document.

3.1. Definitions

Imperial sponsored	Sponsored by Imperial College London (ICL) or Imperial College Healthcare NHS Trust (ICHNT).
Database lock	The process whereby a dataset is ready for analysis and then its state is kept constant – i.e. locked so that the data cannot be subsequently amended.
Interim lock	This refers to the processes used to take a 'snapshot' of a database at a particular point in time while the study is in progress. It involves the temporary locking of the database e.g. for interim analysis, DMC/ DMEC/ Safety review etc. This temporary locking is to prevent trial data being entered or changed whilst RDEs/ data reports are being compiled for the Statistician. When the interim lock is applied, all users with access to add, amend, delete data or queries will have their accounts temporarily disabled or modified to only allow read only access for the duration of the interim database lock. This interim database lock can be reversed, and multiple interim database lock requests can be made and performed.
Final database lock (This is a two-stage	Soft lock:

process, which includes soft database lock and hard database lock)	It refers to processes during which access to the database is limited whilst the suitability of the data for final analysis is determined. The same procedures as a Hard lock are followed, but it is expected that some further activity may be undertaken that may result in data changes. Users' permissions may or may not be restricted during a soft lock. Soft lock should only be requested once all the necessary data management activities have been completed as per the data management and monitoring plan. Hard lock: Refers to the process whereby a clinical trial database has data cleaned and validated and all edit permissions are revoked. Data in a 'hard locked' database is considered clean, complete (as far as is possible) and ready for analysis, and no further data amendments are expected. Users' edit permissions are revoked. The hard database lock for the final database lock should only be applied once and should not be reversed, except in exceptional circumstances and only with agreement from the Sponsor, CI and other senior management if applicable.
Clinical trial database	A repository of data associated with a clinical trial.
Dataset	Refers to a specific data set contained in a specific clinical trial database.
Read Only accounts	A read only account provides users with view only access to data previously entered into the database. A read only user cannot change the data entered into the database. Read only users may be allowed to run reports against data entered into the study, along with any other specific viewing restrictions.
De-activation of accounts	De-activating of accounts is a temporary suspension of an account and can be reversed. This is mainly used for interim and soft database locks
Termination of accounts	Terminating accounts cannot be reversed or re-used.

3.2. Abbreviations

CTIMP	Clinical Trial of an Investigational Medicinal Product
CI	Chief Investigator
CRF	Case Report Form
eCRF	Electronic Case Report Form
EDC	Electronic data capture
DMC	Data Monitoring Committee
DMEC	Data Monitoring Ethics Committee
DMP	Data Management Plan
IA	Interim Analysis
ICHNT	Imperial College Healthcare NHS Trust
PI	Principle Investigator
REDCap	Research Electronic Data Capture

RDE	Research / Report Data Extract
SOP	Standard Operating Procedure
TMF	Trial Master File

4. METHOD

The processes are primarily the responsibility of the Trial management team, although the CI, Statistician, Coordinator, and the Sponsor's representatives may all contribute to the decision to soft and hard lock a trial database.

In order to maintain a robust dataset, it is vital that database locking is a controlled process subject to approvals by the appropriate parties. This is ensured using an appropriate approval process (i.e. using request forms). Please note that a database lock has two parts: 1) the soft lock and 2) hard lock. Both are necessary to provide the final locked database and the associated dataset.

4.1. Data Management procedures for Database Lock

Data management tasks to complete prior to database lock include ensuring:
- All CRF (paper or electronic) data has been received, entered into the database and processed
- All queries have been resolved to the best ability of those raising the queries and those responding to their queries
- External data completed and reconciled (where applicable)
- Safety data completed and reconciled (where applicable)
- Medical coding reviewed and approved (where applicable)
- A final quality control review has been completed successfully

Source document verification of key data must also be completed in time for final data cleaning and query resolution.

4.2. Study specific parameters

A database is closed at the point where all reasonable attempts have been made to collect all outstanding data items/data queries and all study specific parameters have been satisfactorily captured/met.

Prior to database lock the Trial Management team and CI or Sponsor should negotiate where applicable the acceptance level for the following parameters:

- An acceptable number of outstanding queries at database lock (may be defined in terms of critical and non-critical queries)
- An acceptable error rate as determined by final quality control checks (where applicable)
- Any other checks to which the Trial Management team and CI or sponsor agree upon, including checks not related to the database e.g. monitoring queries

If any of the above are not achieved or are not likely to be achieved by the agreed date for database lock, guidance should be sought from the CI or sponsor.

4.3. External Data

If external data importation or direct entry is required for a study the Trial Manager or delegate will ensure the data transmission occurs in sufficient time to allow for data reconciliation and query resolution prior to final database lock. The timelines/intervals for the receipt of external data will have been agreed prior with the external vendor/data supplier.

IMPERIAL

Research Governance
and Integrity Team

NHS
Imperial College Healthcare
NHS Trust

4.4.   Safety Data Review

The Trial Manager or delegate will perform a final safety data reconciliation prior to database lock. All queries raised during the review should be resolved prior to database lock.

If coding is required on a study all coding and approval of medical terms and query resolution must be completed prior to database lock.

4.5.   Process for database lock

A set criterion must be met prior to database lock. This could be documented in the form of a checklist or other documentation, with a full list of procedures to be completed prior to database lock to meet the database lock deadlines.

The EDC software being used should usually allow for two different methods of database case/form locking. Either all forms/cases can be batched locked at the same time or case/forms locked page by page. The method to be used in a study should be described in the Data Management Plan (DMP).

The PI at each site will normally be required to electronically sign off every case form in the study database. The investigators, or appropriately delegated personnel are responsible for reviewing each subject's CRF to confirm the data entered are complete and accurate before database hard lock.

4.5.1 Initiating an interim lock

A member of the Trial team will make the request for an interim database lock under instruction from the CI. The request should be documented in the TMF and ultimately be instigated by the CI. An application is made via the appropriate request system to instigate the locking procedure and provide authorisation to a cleaned dataset. During an interim database lock the Trial team edit permissions are usually revoked.

4.5.2 Initiating a hard lock

A member of the Trial team will make the request for a database lock under instruction from the CI. The request should be documented in the TMF and ultimately be instigated by the CI. An application is made via the appropriate request system to instigate the locking procedure and provide authorisation to deliver a cleaned dataset.

4.5.2.1 Locking the database

Upon receipt of a lock request, authorised Trial team members responsible will begin the process of locking the database. A hard lock is always preceded by a soft lock and should follow in a timely manner.

4.5.2.2 Soft lock / data cleaning

During soft lock, the Trial team will perform final data cleaning activities and ensure it is as complete as possible and contains accurate data. Once the locking process has been initiated most of the data should be in place and routine addition of new data is not expected. However the addition of missing data as directed by the Statistician is permissible. The dataset may or may not require further revisions, the Statistician may highlight areas that need further cleaning and/or missing data that are required. Data changes during this

process should be undertaken in the same way as routine data management and data entry, this is an iterative process until the Statistician is satisfied. During soft lock, it may be the case that edit permissions are retained in order that data can be updated. It is permissible however to remove permissions from users who no longer require edit permissions e.g. for a user from a site that has had all data cleaned.

4.5.2.3 Hard lock

Only when soft lock is completed, and the dataset is considered final and ready for analysis can the hard lock be initiated. Evidence of procedures and processes that led to the decision that the dataset is final must be retained. This includes any correspondence, reports, checklists and assessments generated during the cleaning process.

Once database hard lock has been completed all user accounts will be terminated with the exception of those identified as having read only access for reporting and review purposes on the study and this will only normally be for a limited period of time.

4.5.3 Unlocking the database

The hard database lock for the final database should only be applied once and should not be reversed, except in exceptional circumstances. Database unlock occurs where the locked database is altered and is made available for further changes. If a discrepancy or query is identified after the database has been locked, then access to the database may need to be made so that changes can be made. Unlocking will be limited to significant corrections that will have an impact on the reliability of the results. In the event that a locked database requires unlocking, the CI should seek permission from the Sponsor. The justification to unlock the database, the effect on the statistical outcome and the completion of the associated approval process must be documented in the TMF prior to the unlock being undertaken.

4.5.4 Relocking the database

Prior to re-locking, the audit trail for the database must be reviewed to confirm that only the approved changes were made. Any files, data etc. created to support this process should be retained. The outcome of the review and if appropriate, the audit trail should be filed in the TMF. If it becomes apparent that data-points not approved by the sponsor have been amended, this must be escalated to the Sponsor, CI and trial statistician immediately and appropriate steps taken to ensure data accuracy, compliance and correct documentation.

Only once all queries or issues have been resolved the re-locking process should be undertaken again as soon as reasonably appropriate.

4.5.5 Locking with unclean data

In some cases, the trial team may not clean every data point. In these cases, the decision to lock the database should be based on the requirements of the protocol (e.g. when the data is not required for the endpoints of the trial or a sufficient number/percentage of endpoints entries shall suffice). The Statistician and the CI must be involved in this discussion and decision making process.

IMPERIAL

Research Governance
and Integrity Team

NHS
Imperial College Healthcare
NHS Trust

Any data that is knowingly left unclean should be documented as such, and this documentation should be available to consumers of the data.

4.6. Extension studies

Extension studies typically follow a double-blind randomised placebo-controlled trial of a CTIMP. At the end of the double-blind phase, participants are invited to enrol in an extension study. The study will normally be longer than the randomised trial with the objective primarily to gather information about safety and tolerability of the new drug in long term, day to day use.

In extension studies if there is a need to lock the data for the main study, whilst still allowing continued data entry for the extension study without risking changing the main study data, this may be approved.

The database lock process for extension studies should be dealt with on a case-by-case basis and detailed in the study DMP. Normally the main study would be granted an extension and therefore the primary end point analysis will be considered an interim analysis (interim lock), and the extension data will be considered the final analysis (hard lock). Once data cleaning of the extension data has been completed the hard lock of the full database can be completed as per section (4.5.2.3) above.

5. REFERENCES

1. MHRA Good Practice 'Grey Guide'
2. REC SOP version 7.6 September 2022
3. Imperial Clinical Trials Unit SOP IN007: Database Lock for InForm Clinical Trials
4. Imperial Clinical Trials Unit SOP: DM015: Database Lock Procedures
5. RGIT_SOP_020 Data Management