


## GD013 Database Locking

Database locking is the process whereby a dataset is made ready for analysis and locked so the data cannot be amended.

There are two types of lock:

1. An **Interim Lock** or **Soft Lock** is a process that provides a snapshot of a database at a particular point in time while the study is still in progress. This is usually only performed if specified in the protocol, the exception would be if agreed by the Study Steering Committee or similar and agreed by the sponsor. The lock is only required until the data has been provided to the statisticians, or similar as appropriate, for the snapshot.
2. A **Hard Lock** or **Database Lock** refers to the process whereby a database has been 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; no further amendments are expected.

The following steps are part of lock preparation. What is required for an interim lock and a hard lock will be documented in the Data Management Plan (DMP) (see GD016). In most cases, steps 1, 2, and 4 are combined in a data completion review. However, the items marked with \* are always required for a hard lock:

1. **Case Report Form (CRF) completion review:** The database should be reviewed to ensure the appropriate CRFs have been completed for each participant.
  - a. If the study is using OpenClinica, the CRFs should be marked completed .
  - b. The exceptions to all CRFs being marked as complete are the use of:
    - i. Skipped – only to be used on admin CRFs or similar forms used by the study team for study management, or when a follow up visit did not take place.
    - ii. Stopped – used when a CRF has been scheduled in error.
    - iii. Unscheduled – an event or CRF will show as unscheduled if it has not been required for use for that particular subject ID, an example of this is where participation has ended early.
  - c. All missing data should be reviewed, where necessary an annotation note can be added to provide a reason for the missing or unexpected data – this would be a requirement for a required data field (denoted by an \*).
2. **Data query review** \*: A final review of the data should be made to identify and resolve any remaining queries. Any unresolved data queries should be documented in a file note and signed off by appropriate parties.
3. **Automatic edit check review:** These are usually run via an automated program, after the data has been entered, generating a report listing the possible validation issues. All validation issues should be reviewed prior to lock. Any unresolved validation issues should be documented in a file note and signed off by the appropriate parties.
4. **Source data verification review** \*: The level of source data verification (SDV) specified in the monitoring plan should be achieved by the study monitor prior to lock. The review independently verifies that this has been completed, and no changes made to the database have unmarked a CRF SDV'd without the monitors knowledge.

5. **Serious adverse event reconciliation:** If required by any stakeholder, any adverse event reported as a serious adverse event in the database should be reconciled against events reported to the stakeholder's pharmacovigilance. If this reconciliation is required by a non-Papworth sponsored study, the sponsor's SOP and guidelines should be followed.
6. **Coding:** For studies where the DMP specified coding; all coding must have been completed prior to a hard lock, for an interim lock this will be specified on FRM053.
7. **PI Signature:** The Principal Investigator (PI) from each site should sign off each patient's casebook as part of their GCP requirement. They are confirming that the CRFs are the full, accurate, and complete record of the visits.
  - a. For studies using OpenClinica, PI should sign off the patient's dataset using the built-in function (for CTIMPs this must happen). When a casebook for a subject is signed, OpenClinica automatically sets the status for all study events for that subject to 'signed'. After an event status is 'signed', any changes to the CRF automatically take back the event status as 'completed'. OpenClinica also allows for the signing of individual events if required.
  - b. If the study is using OpenClinica and is non-interventional then the CRFs can be signed off on paper or by using an e-signature system like DocuSign by the Chief Investigator. This process must be approved by the Sponsor.
  - c. If the study is not using OpenClinica the CRFs can be signed off on paper or using an e-signature system like DocuSign.
  - d. If any data is changed following PI sign off, they will require the PI sign off again.
8. **Authorisation to lock obtained from all parties \*:**
  - a. If an interim lock is required, FRM053 'Interim Database Lock Form' must be completed specifying what data, subjects and/or CRFs are to be included. This form should also be used to record which steps have been completed prior to interim lock.
  - b. If a hard lock is required FRM017 'Database Lock Form' must be completed. FRM017 must also be used to record any steps have been completed prior to hard lock. See below for more details on this form.

Once the steps specified in the DMP have been completed, the data manager may run a final check on all primary and secondary endpoints before locking the database. The method for locking will depend on the data repository used.

For a hard lock, the database should have all edit permissions revoked to prevent any amendments to the database. If not using OpenClinica as the data repository, how the lock will be achieved should be documented in the DMP, especially for any interim lock. As getting the snapshot of the data without any data changing may be harder in some systems, and extra steps may need to be taken.

After successfully locking the database, the data manager should notify the relevant stakeholders.

### Unlocking the database

1. **Hard Locked Database.** The dataset should not be amended after a hard lock, but it can be unlocked if necessary. Unlocking a hard locked dataset should be avoided as this can only be approved through a formal process at sponsor level. The following steps are always required when unlocking and relocking a database:
  - a. Complete FRM019 'Database Unlock Request Form' to obtain the chef investigator, the statistician and the sponsor's authorisation to unlock.
  - b. Prior to re-locking, the audit trail for the database should be reviewed to confirm that only the approved changes were made. Any files or data created to support this process should be retained.
  - c. A new FRM017 'Database Lock Form' should be used to authorise the relock.
  
2. **Interim Locked Database.** Data that is locked in an interim lock can be unlocked at a timepoint after the lock, without any additional approvals or forms. Before each interim lock, it will be decided if this will happen, and when. For example, it might be determined that all the data will be unlocked, but time will be allowed to extract the data for the lock. Or it might be decided that the data should remain locked until final database lock, with the caveat that if a patient or visit/event needs unlocking for cleaning or monitoring that this can happen. Or it can be something in between. Whichever it is it must be clearly recorded on the FRM053 'Interim Database Lock Form' which participant datasets will be unlocked, and which will remain locked following the interim analysis. Best practice would have somewhere, like a spreadsheet, recording both the original lock, and any unlock, and this should be noted on FRM053.

Referenced Documents
GD016 Data Management Plan
FRM019 Database Unlock Request Form
FRM017 Database Lock Request Form
FRM053 Interim Database Lock Form