UTS IMPACCT Trials Coordination Centre

Standard Operating Procedure

5.5.2 Electronic Data Transfer

Version	V2.5
Author/s	J Hao
Approved	M Agar
Effective date	01/04/2022
Review date	01/04/2024

DO NOT USE THIS SOP IN PRINTED FORM WITHOUT FIRST CHECKING IT IS THE LATEST VERSION AS AVAILABLE FROM www.uts.edu.au/itcc

Introduction / Background

Data for clinical studies may need to be stored, checked, and managed in a separate location from where the data will be monitored or analysed. ITCC/PaCCSC/CST need to ensure that study data are transferred from one person and location to another in a timely, safe, and secure manner that will maintain data integrity and participant anonymity.

Objective

This SOP describes the procedure for data transfer for data generated within the IMPACCT Trials Coordination Centre (ITCC) including the Palliative Care Clinical Studies Collaborative (PaCCSC) and Cancer Symptom Trials (CST). In most cases this data transfer occurs when the study statistician needs to check allocation codes against study data (for data safety monitoring) or for study analysis (interim or final). Subject to approval by the Coordinating Principal Investigator, others may request access to the study data at various times throughout the study.

Scope

This SOP applies to all sites involved in clinical studies conducted by ITCC/PaCCSC/CST. It also applies to all staff involved in clinical studies irrespective of individual organisational employment, role or position.

Ownership and Responsibility

It is the responsibility of the Principal Investigator to ensure that all data and files are accurately managed. The task may be delegated to another suitably trained individual as documented on the Staff Signature and Delegation Log (*refer* SOP 4.2.4 Delegation of Duties) but the responsibility remains with the Principal Investigator.

Responsibilities of the ITCC National Project Officer

- To download data from the Coordinating site and save on a secure drive as a back up
- To save the data with clear version control to enable changes and updates to be tracked over time

Procedure

1. Data download

This procedure is undertaken by the ITCC National Project Officer at the Coordinating site, and downloaded onto a password protected network drive, with clear and secure backup and retrieval procedures.

On each occurrence, the following procedure is followed:

- Data download is saved to password protected computer as an excel workbook or CSV text file with file name indicating Case Report Form (or other data file) name and date.
- The downloaded file is viewed for completeness and gross validation. This is not data checking (*refer* SOP 5.5.1 Electronic Data Handling) but a check to see that all data fields have been downloaded in a complete and consistent manner.

2. Database checking

The database for each study is checked by the Coordinating site at the time of download for the following elements:

- To identify and remove any duplicate or blank entries
- Correct any participant ID numbers that have been incorrectly entered, and where the ID number can be verified from another source (for e.g., email confirmation or Randomisation Registration Notification)

File notes and data report forms are used to document changes to the data base.

3. Data safety

This procedure is undertaken when there has been an authorised specific request for data from an external source (*refer* SOP 5.5.10 Data Ownership and Utilisation):

- Each data file is protected within Excel using encryption and password through the Protect Worksheet function.
- All files are then “zipped” into a WinZip file. The data file is sent via email to the responsible party (this may be the ITCC, Coordinating Principal Investigator, Data Safety Monitoring Committee member or Lead Statistician depending on the data and the request).
- The password to open the zipped data file is forwarded to the responsible party separately via telephone or by email where the subject line does not link the two emails.
- The responsible party opens the encrypted data file using the password and imports the data into an appropriate statistics programme using the import function.
- Field and data codes are entered as per the study data dictionary in order to facilitate analysis.
- The main study data are merged with the table of allocation codes by the responsible party. This will only take place in the following situations:

- For unmasked data analysis as requested by the relevant Data Safety Monitoring Committee
- At the completion of study accrual to facilitate all study analyses

4. Database closure

- Database closure occurs when:
  - Study recruitment has been stopped
  - All data has been entered and checked
  - The data meets the definition of Clean Data
- The database for each study is checked by the ITCC at the time of download for the following elements:
  - To identify and remove any duplicate or blank entries
  - Correct any participant ID numbers that have been incorrectly entered, and where the ID number can be verified from another source (for e.g., email confirmation or Randomisation Registration Notification)
- Changes to the database are recorded within the system audit trail.
- After a proper quality check and assurance, the final data validation will be run. If there are no discrepancies, the datasets will be finalised in consultation with the study statisticians.
- All data management activities should have been completed prior to database lock. To ensure this, a pre-lock checklist (Template 50) will be used, and the completion of all activities will be confirmed. This is done as the database cannot be changed in any manner after locking. Once the approval for locking is obtained from all stakeholders, the database will be locked, and clean data will be extracted for statistical analysis. Generally, no modification in the database is possible. But in case of a critical issue or for other important operational reasons, privileged users will be able modify the data even after the database is locked. This, however, requires proper documentation and an audit trail has to be maintained with sufficient justification for updating the locked database.
- Data extraction will be done from the final database after locking. This will be followed by its archival.
- The main roles acting in this phase are:
  - **Data entry staff:** responsible for updating of the database before the final closure analysis;
  - **Project Data Manager**: responsible for completing database closure checks and locking of data in REDCap EDC system and suspending of the access rights to the database.
  - **Project Manager**: responsible for approving the database lock and filing the completed approval form in Trial Master File (TMF).

5. Closure checks

Closure checks refer to the verifications that should be performed prior to database lock to verify the integrity and completion of the study database. They will include:

- Check that all expected eCRFs have been entered;
- Check that all the queries are resolved;
- Check that the database is consistent;
- Determine the status of each participant entered (i.e., excluded, ongoing, completed, withdrawn, lost to follow-up, etc.);
- Check for value formatting problems in database exports;
- Confirm that all expected site signatures have been applied.

The data closure checklist is signed and filed in the TMF.

6. Database Lock and Preparation for Final Analysis

- After the date of database closure, all the data are downloaded and filed in a password protected network drive, with clear and secure backup and retrieval procedures. All the records are locked in the REDCap EDC system. This point is considered Database Lock. Database lock is the time point, for a clinical trial, at which a database is expected to be clean, all data are completed and consistent, all queries resolved, and a final quality control has been performed, so that the database is ready for final analysis. Once all data has been transferred to the study statistician and any resultant queries have been resolved, the database is locked, all continuing access is ceased, except for the data manager or coordinator.
- This process will follow these steps:
    - Authorisation for removal of the access rights to the REDCap EDC system using the Data Closure Checklist and Lock Approval Form (Template 50).
    - Removal of access rights to the database done by the database administrator or data manager. Export of data for the analyses. This ‘final’ database should be clearly indicated with the date (and time if applicable). In addition, exported data files are preferable saved as read-only files.
- The data manager should document the status of the database on the Data Closure Checklist and Lock Approval Form (Template 50). The project is responsible for the approval of the database lock recorded and documented the approval form in TMF.
- An audit should be performed of a sample of completed forms in the REDCap EDC system against exported datasets to ensure the integrity of the final study data.

7. Updating the Database after Database Lock

- Updates to a locked database should be limited to important corrections, for instance if the data to be changed have a significant impact on the reliability of the results.

- If the statistician has data queries following database lock, they will need to provide justification for requesting the unlocking to the REDCap EDC system provider and detailed reason for changing the data.
- Both request and approval will need to be documented on a Database Unlock Request Form (Template 49).
- Only the designated data entry staff member (as detailed on the study delegation log) will be granted access to the database again to implement the required changes.
- The database will be re-locked as soon as the corrections have been made to prevent other data changes. Once the database is locked again, a new final database file (not overwriting the original database lock) will be created.
- This process will not only be recorded on a Data Closure Checklist and Lock Approval Form (Template 50) by the ITCC Project Data Manager, but also approved and filed in TMF (*refer* SOP 5.5.3 EDC access). Re-locking of the database must be performed according to procedures set out in section 6.
- An audit trail will be implemented, listing the database lock date, and keeping details of what has been changed and when.

8. Database transfer to CPI/ study statistician

- Dataset Specifications: File Transfer Format
- UTS-ITCC will provide a final locked database to the Coordinating Principal Investigator (CPI) or study statistician upon request.
- All files “zipped” into a WinZip file are sent via OneDrive (for business) to CPI/ study statistician. "OneDrive for Business" with its file-sharing security features and encryption (at rest and in-transit), is suitable for transferring the UTS highest classification - "UTS Confidential" data.
- The UTS-ITCC team sign in with their Microsoft account or work account to access OneDrive (for business).
- UTS-ITCC team ‘share’ the files which will insert a link to the OneDrive file in an email message sent to the CPI/ study statistician. Access to the link is granted to the appropriate individuals only and the link only will work until the date is set. The expiration date will be set for 24 hours after sending the link. Once it’s expired, the link will be invalid.
- The password to open the zipped data file is forwarded to the CPI/ study statistician separately via telephone or by email where the email subject line does not link the two emails. When the CPI/ study statistician clicks the link, they will be prompted to enter a password before they can access the file.
- Once the encrypted data file is opened using the password, the CPI/ study statistician imports the data into an appropriate statistics programme using the import function.

- The format for data transfer will be using Excel files. CSV files can be converted to various other formats using commercially available off the shelf software. Each dataset will be provided in a single Excel file, or those datasets which are divided will be clearly named to aid the reviewer in reconstructing the original dataset. The transfer document will identify the range of participant numbers (or other criteria used for division) in the label for each of the divided datasets.
- For all datasets, in order to significantly reduce dataset file sizes, the allotted character column length/size for each column will be the maximum length used. All dataset names and dataset labels are unique across the tabulation datasets submitted for the study. The internal name for the locked dataset on these CSV files will be the same as the variable/field name shown in the data dictionary codebook. The key variables will appear first in the datasets. Each participant will be identified by a single and unique participant identifier code.

Related SOPs

4.2.4 Delegation of Duties

5.5.1 Electronic Data Handling

5.5.3 EDC access

5.5.10 Data Ownership and Utilisation

8.0 Essential Documents

Related documents

Template 50: Data Closure Checklist (all sites) and Lock Approval Form

Template 49: Database Unlock Request Form

References

Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95). Annotated with TGA comments 2000 (accessed 07/02/2020)

https://www.tga.gov.au/sites/default/files/ich13595an.pdf

History			
Version	Date	Author	Reason
1.1	18/07/2007	B Fazekas	New procedure
1.2	18/08/2007	B Fazekas	Changes ratified by MAB
1.3	16/10/2007	B Fazekas	Update after David Currow review
1.4	9/06/2010	B Fazekas	Periodic review
2.0	3/02/2011	B Fazekas	Changes ratified by MAB
2.1	30/12/2015	B Fazekas	Periodic review
2.2	28/02/2018	B Fazekas, S Kochovska	Periodic review Publication of the ICH GCP E6 (R2)
2.3	06/12/2018	L Brown	Update to include CST and ©
2.4	07/02/2020	J Lourdesamy	Periodic review
2.5	07/12/2021	J Hao	Periodic review

Approval		
Version	Approval Name	Approval Signature
2.5	M Agar