Standard Operating Procedure
Research Governance

Title:	**Data Management: Collection, Validation and Storage**		
SOP Reference Number:	QUB-RGEI-014	Version Number:	FINAL v 1.0
Revision Date:	20 September 2021	Review Date:	20 September 2024

	Name and Position	**Signature**	**Date**
Author:	Research Governance, Ethics and Integrity Team	----------	----------
Reviewed and Approved by:	Chair, Research Governance, Ethics and Integrity Committee	----------	----------

This is a controlled document.
When using this document please ensure that the version is the most up to date by checking the Research Governance, Ethics and Integrity Website

Revision Log

Previous Version number	Date of Review/Modification	Reason for Review/Modification	New Version Number

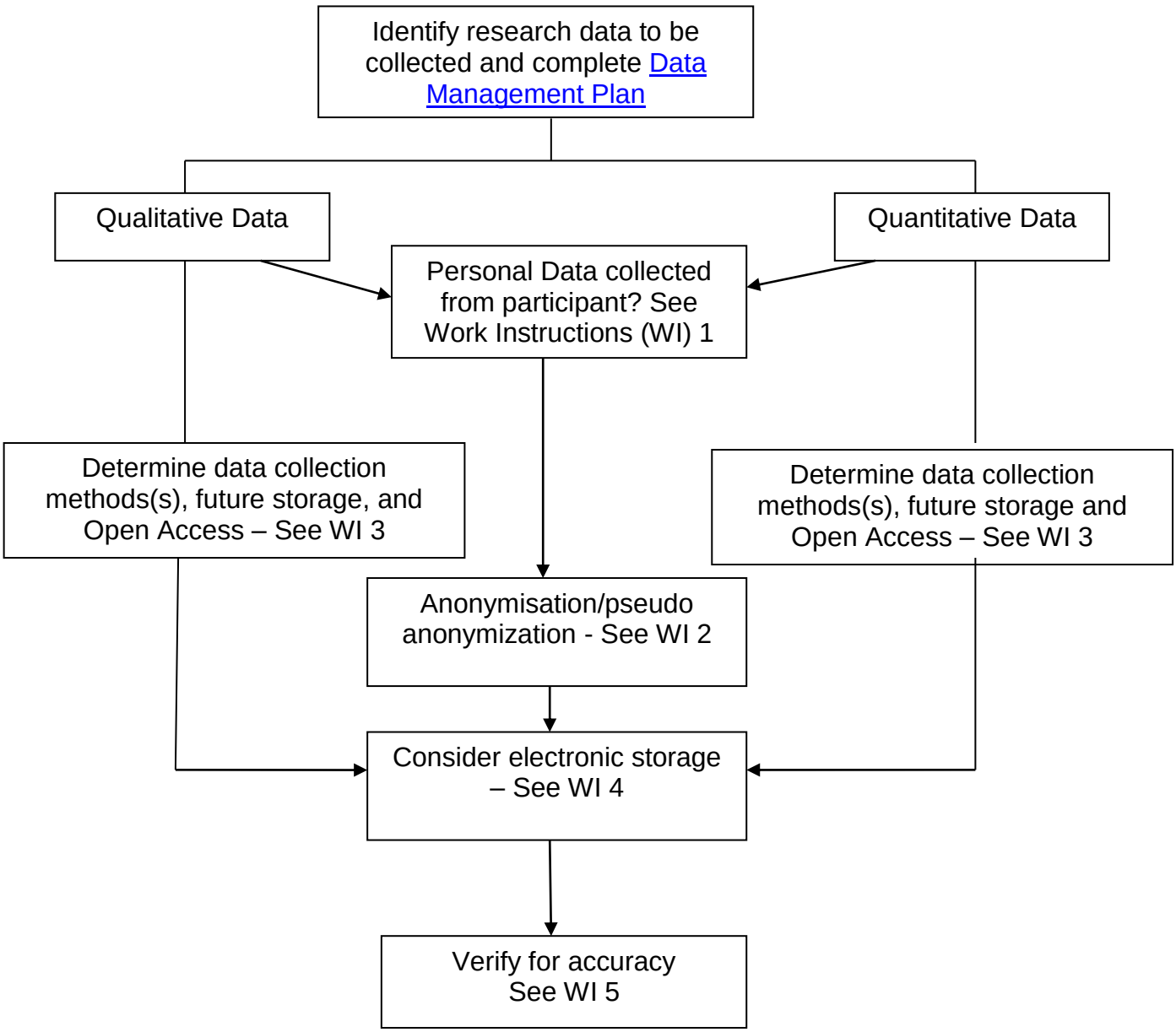
1. Purpose

This Standard Operating Procedure (SOP) describes the procedure for the management of research data and ensuring that active research data is held in a safe and secure manner.

2. Scope

This SOP applies to all studies where the University is acting in the capacity of Sponsor, or Co-Sponsor. It applies to all members of University staff; both academic and support staff as defined by Statute 1, including honorary staff and students.

3. Procedure



4. References

Data Protection Act 2018.
https://www.gov.uk/government/collections/data-protection-act-2018
(last accessed September 2021)

Queen's University Belfast, Guide to Handling Personal and Sensitive Data - Research
https://www.qub.ac.uk/about/Leadership-and-structure/Registrars-Office/FileStore/Filetoupload,945308,en.pdf
(last accessed September 2021)

Queen's University Belfast, Information Services, Information Security
https://www.qub.ac.uk/directorates/InformationServices/Services/Security/ (last accessed September 2021)

Data Management Team, Library Services
https://libguides.qub.ac.uk/ResearchDataManagement

5. Appendices

Work Instructions 1 – Collection of Personal Data
Work Instruction 2 – Anonymisation/Pseudoanonymisation
Work Instruction 3 – Data Management of Physical Records
Work Instruction 4 – Electronic Data Management
Work Instruction 5 – Data Validation

Work Instructions 1 – Collection of Personal Data

1. All planned research that intends to collect personal data must first take cognisance of the University's "Guide to Handling Personal and Sensitive Data".
2. Personal data is considered as information that relates to an identified or identifiable individual, for example a name, location data, date of birth, on-line identifier such as IP address.
3. Sensitive data includes race, ethnic origin, political beliefs, religious or philosophical beliefs, trade union membership, genetic data, biometric data, health data, sex life or sexual orientation.
4. Where personal or sensitive data is to be collected as part of research, the researcher must comply (General Data Protection Regulations) GDPR requirements, outlined in the University's guide, referenced above.

Work Instructions 2 – Anonymisation/Pseudoanonymisation

1. Pseudoanonymisation may involve the replacement of names or other identifiers with a reference number.
2. The reference number is kept separate from the research data, but it still provides the key to unlock participant's identification.
3. Pseudoanonymised data remains personal data and within the scope of GDPR.
4. Anonymisation is when personal data cannot be traced back to an individual and there is no key to link back. Anonymised personal data is not subject to GDPR.
5. Where confidential personal data (e.g. names, addresses, dates of birth) must be retained, this should be held separately and access limited to designated persons.
6. Anonymisation/Pseudoanonymisation can be achieved through coding by being given a unique identifier by the investigator. This should be used on the electronic data management system to allow for the identification of all the data reported for each research participant.

Work Instructions 3 – Data Management of Physical Records

1. Consider the quantity of storage required to maintain physical records – this may necessitate off-site storage.
2. If using off-site storage ensure there are appropriate agreements in place
3. The security of any storage must be carefully considered to ensure the research record and/or personal information is not compromised.
4. The University requires research data to be stored from five years from completion of the study, ensure Funder's Terms and Conditions are met if these stipulate a longer requirement.
5. Ensure the research records are captured as part of the School's data retention scheme so the records can be disposed of at the end of the retention period.
6. Consider requirements for Open Access and discuss with the Data Management Team in Library Services.
7. Determine how the records are be disposed of – do they require shredding/confidential waste management.

Work Instructions 4 – Electronic Data Management

1. Whether the research data is qualitative or quantitative paper records are often transferred onto electronic sources to support data management.
2. Qualitative data may be the transcription of interviews, quantitative may be the establishment of data base used to capture study design in accordance with the study protocol.
3. It is important to ensure that electronic data management enables:
    a. Completeness;
    b. Accuracy;

   c. Reliability;
   d. Consistency.
4. To facilitate electronic data management, a Standard Operating Procedures for the use of any database should be developed.
5. Database must be secure and unauthorised access to the data prevented through the use of appropriate password-protection.
6. Rather than data being deleted if changes are required, changes should be captured.
7. Adequate backup of the data must be maintained. Storage of data on the University's server provides automatic back-up. If it is not possible to use the University's server, data must be backed up regularly and the back-up files kept securely in a separate location of the master copy.
8. The University requires research data to be stored from five years from completion of the study, ensure Funder's Terms and Conditions are met if these stipulate a longer requirement.
9. Participant data, being entered onto the study database, should be anonymised at the earliest opportunity.
10. Any data held on portable equipment such as laptops, CDs, DVDs, memory sticks or portable hard-drives must be risk assessed in accordance with the University's Mobile Computing Policy and this risk assessment must take into account the sensitivity of the information. In addition, portable equipment must have password and time-out protection and encrypted.
    Further information can be accessed on the information services website:
    https://www.qub.ac.uk/directorates/InformationServices/Services/Security/
11. Ensure that the person(s) responsible for data entry has received the necessary training to undertake this task.
12. Prior to data entry incomplete or missing information, and any inconsistencies should be checked with the researcher.
13. Coding data entries in quantitative research allows data to be transferred on the study database. Consideration should be given as to how codes are to be determined and used consistently. It is important to ensure a value is provided for missing data, 'not known' or 'not applicable' so no field is left blank.

Work Instruction 5 - Data Validation

1. It is important to ensure that the data entered into the study data management database is accurate. Therefore it is necessary to validate the data and 'clean' it when necessary.
2. This can be achieved through a variety of methods and depending on the software package used to develop the database. A SOP should be prepared to ensure that data is validated in a consistent manner.
3. At the end of each data entry session, the person responsible for data entry should check their entries for any obvious errors looking for missing values or values outside of a pre-defined range.
4. If employed on a study, the data manager should also validate the data through the creation of reports and/or lists and looking for logical discrepancies.
5. Data validation should be continued until all missing values and inconsistencies are corrected and clarified.