Queen's University
Belfast

Research and Enterprise

Standard Operating Procedure
Research Governance

Title:	**Data Management: Collection, Validation and Storage**		
SOP Reference Number:	QUB-ADRE-017	Date prepared	6 August 2008
Version Number:	Final v 5.0	Revision Date	19 Jan 2017
Effective Date:	01 March 2017	Review Date:	January 2019

	Name and Position	Signature	Date
Author:	Mrs Louise Dunlop Head of Research Governance	Louise Dunlop	29-03-2017
Reviewed by:	Professor James McElnay, Chair Research Governance and Integrity Committee	James McElnay	22-03-2017
Approved by:	Mr Scott Rutherford Director, Research and Enterprise		15.3.2017

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 Website

* For all University sponsored research recorded as risk category level 4, including IMP studies
For all other University sponsored research involving human participants

Revision Log

Previous Version number	Date of Review/Modification	Reason for Review/Modification	New Version Number
Final v 1.0	10/11/2009	Annual Review	Final v 1.0
Final v 1.0	09/09/2011	Annual Review/ Update following MHRA GCP Inspection	Final v 2.0
Final v 2.0	21/08/2012	Periodic Review	Final v 3.0
Final v 3.0	23/10/2014	Periodic Review	Final v 4.0
Final v 4.0	19/01?2017	Periodic Review	Final v 5.0

1. Purpose

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

The primary focus of this SOP is for clinical trials of Investigational Medicinal Products (IMPs) but it is relevant for any research being undertaken under the auspices of the University.

2. Introduction

The International Conference on Harmonisation (ICH) Good Clinical Practice (GCP) guideline specifies that *"appropriately qualified individuals supervise the overall conduct of the trial, handle the data, verify the data, conduct the statistical analyses and prepare the trial reports" (ICH 5.5).* A requirement of GCP is that "*the confidentiality of records that could identify subjects be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirements*".

It is therefore necessary that only essential clinical trial data for the purpose of the study is collected, managed in accordance with the GCP requirements and stored in accordance with the Data Protection Act (DPA) 1998. The DPA stipulates eight principles that must be adhered to.

These are, in summary:

- That personal data must be fairly and lawfully processed;
- Processed for limited purposes;
- Adequate, relevant and not excessive;
- Accurate and up to date;
- Not kept longer than is necessary;
- Processed in accordance with the rights of the data subject;
- Secure;
- Not transferred to countries outside the European Economic Area (EEA) without adequate protection.

Researchers storing or handling personal data must comply with the DPA.

Data management involves the translation of data captured on the Case Report Form (CRF) into electronic data for the purpose of statistical analysis.

3. 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.

4. Responsibilities

4.1 Chief Investigator

The Chief Investigator (CI) is responsible for ensuring that the research participant is fully informed as to who will have access to the participants data, where it will be stored, by whom, how and how the person's data will be processed. This is most appropriately undertaken through the informed consent process. The CI is also responsible for ensuring that research data is captured, processed and handled correctly.

4.2 Site Principal Investigator

The Site Principal Investigator (SPI) on a local research site is responsible for ensuring that research data captured locally is processed and handled correctly in accordance with the requirements of the site organisation and this SOP.

5. Procedure

5.1 Data Collection

All requests to collect personal information should clearly state:

- The purpose for which the information is to be used;
- The period of time it is to be retained;
- To whom it is likely to be disclosed.

Sensitive personal data should only be processed if participants have given their informed consent and it has been highlighted to them that it will be stored on University premises (where a Health and Social care patient/client is the research subject). It should also be highlighted that the individual's data may be viewed by members of the research team, regulatory authorities or University/Trust research governance staff.

The case report forms should be completed and stored in accordance with SOP QUB-ADRE-007.

5.2 Electronic Data Management

The database that the CI uses to capture the study data should be designed in accordance with the protocol and to encapsulate the information collected on the CRF. When this database is being designed it is important to ensure that it meets the ICH GCP requirements for:

- Completeness;
- Accuracy;
- Reliability;
- Consistency.

The ICH GCP guidance states that SOPs for using the database should be in place. These should be study specific and developed by the CI or a designated individual. If a third party is being used to design, build and maintain a study database there should be a formal contract requiring compliance with ICH GCP and regulatory expectations.

The database should be established in a way that changes are permitted but that data changes are documented. There should be no deletion of entered data.

The database must be secure and unauthorised access to the data must be prevented through the use of appropriate password-protection.

A list of individuals who are authorised to make data changes should be maintained.

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 should be backed up regularly. The back-up files must be kept securely in a separate location of the master copy.

Participant data, being entered onto the study database, should be anonymised at the earliest opportunity. This can be achieved through coding (see paragraph 5.3).

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.

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: http://www.qub.ac.uk/directorates/InformationServices/Services/Security/FileStore/Filetoupload,294891,en.pdf

Finally, ICH GCP requires that if blinding is involved in the study, the data entry and processing systems allow for the blinding to be maintained.

If data is transformed during processing, it should always be possible to compare the original data and observations with the processed data.

5.3 Coding

Each subject should be anonymised by being given a unique identifier by the investigator. This should be used on the database as it allows for the identification of all the data reported for each subject.

Responses on the case report forms should be coded, either numerically or alphabetically, to allow for the data to be transferred on to the database. The codes should be determined from the outset and used consistently. It is also important to ensure a value is provided for missing data, 'not known' or 'not applicable'. Fields should not be left blank.

5.4 Data Entry

It is necessary to ensure that the person(s) responsible for data entry has received the necessary training to undertake this task.

Prior to data entry each case report forms should be checked for incomplete or missing information and any inconsistencies checked with the researcher. A record should be kept both the query raised and the response received.

Once content the case report forms is complete, the person responsible for data input should enter the data accurately and the case report forms stored in accordance with SOP QUB-ADRE-007.

5.5 Data Validation

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. 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.

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.
The data manager should also validate the data through the creation of reports and/or lists and looking for logical discrepancies.

Data validation should be continued until all missing values and inconsistencies are corrected and clarified.

5.6 Data Protection

Compliance with the Data Protection Act 1998 is paramount and the participants' confidentiality must be maintained at all times.

Once data has been anonymised it is necessary to hold the key to the participants' code separate to their case report form. This key should be held in a secure location.

Care must be taken when transferring case report forms to other sites, either as part of a multi-centred study or to facilitate the practicalities of a single site study, to prevent against loss or damage.

In the event of electronic data transfer, this should be via a secure system, password protected and encrypted where possible. The University's Information Security Policy provides detailed information on this process.
http://www.qub.ac.uk/directorates/InformationServices/Services/Security/FileStore/Filetoupload,606340,en.pdf (last accessed 19 Jan 2017)
Personal information, no longer required, should be disposed of in a manner that protects its security and confidentiality.

6. References

International Conference on Harmonisation (ICH) of Good Clinical Practice (GCP).
http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html
(last accessed 19 Jan 2017)

Data Protection Act 1998.
http://www.legislation.gov.uk/ukpga/1998/29/contents (last accessed19 Jan 2017)

Queen's University Belfast, Research Data Management Policy
https://home.qol.qub.ac.uk/webresources/_layouts/15/WopiFrame2.aspx?sourcedoc=/webresources/Research%20and%20Enterprise/All%20Qub/QUB%20Data%20Management%20policy%20Feb%202015.docx&action=default
(last accessed Jan 2017)

Queen's University Belfast, Information Services, Information Security Resources

http://www.qub.ac.uk/directorates/InformationServices/Services/Security/#Info
(last accessed 19 Jan 2017)