STU-SOP-DMS-012 – Standard Operating Procedure on Sharing Participant Research Data

1. Purpose and definitions

This SOP applies to research data held on Swansea Trials Unit (STU) servers and managed by STU. It is in line with legislation regarding how data is handled and shared including the GDPR which is implemented in the UK by the DPA, and the Good Clinical Practice (GCP) guidelines. This SOP is in alignment with Swansea University's Data Protection Policy and the UK Policy Framework for Health and Social Care Research (2017).

This SOP does not apply to individuals or trial teams who access data through a local or national Data Safe Haven (aka Trusted Research Environment (TRE)).

The SOP will be used as the basis for any specific data sharing agreement.

Definitions	
Data Controller	An organisation or person who (alone or in common with others) determines the purposes and means for any personal data to be processed. They act on their own autonomy.
Data Processer	An organisation or person who processes personal data under the instruction of the data controller.
Data Requester	An organisation or person who requests access to data held on data subjects in accordance with applicable legislation.
Data Subject	A living person who can be identified, directly or indirectly through information being held that relates to and is specific to them as an individual.
Data Protection Officer	Sometimes called the Information Compliance Officer, is responsible for ensuring employees of an organisation are aware of, trained on and comply with the UK General Data Protection Regulation (GDPR).
Information Commissioners Office	The UK's independent authority set up to uphold information rights in the public interest, promote openness by public bodies and data privacy for individuals. They deal with the Data Protection Act (DPA) 2018, GDPR (2016) and the Freedom of Information (FoI) Act (2000) amongst other regulations.
Personal data	Any information that relates to an identified or identifiable individual. It includes confidential data which is not in the public domain and sensitive data as defined by the DPA.
Pseudonymised data	The processing of personal data by technical and organisational methods in a manner that ensures personal data are not attributed to any data subject. This is a security measure which reduces the risks to data subjects and helps data controllers meet their data protection obligations.
Anonymised data	Where personal data has been stripped of sufficient elements and the individual will no longer be identified or re-identified using any reasonable means. Truly anonymised data is not subject to the UK GDPR.

2. Background

Researchers are expected to maintain high ethical and governance standards as required by employing and funding organisations, professional bodies and legislation when collecting personal data of data subjects. Wherever possible data is anonymised in research, however, pseudonymised data with limited members of the research team able to access identifiable data is also common. Such data can be shared legally and ethically with appropriate safeguards in place .

The lawful basis on which to share and process data is UK GDPR and Data Protection Act (2018). While best practice is to obtain consent, non-identifiable data can be shared without consent of the data subject. The GDPR describes a legitimate purpose for processing of personal data as "*necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller*". Therefore, any data shared will be expected to fulfil this requirement. Requesters are contracted to use the data to generate new knowledge and understanding with the intention to publish their research findings for the wider scientific community and eventual public benefit.

Data will only be shared with suitable organisations/individuals with a Data Sharing Agreement (DSA) or Data Access Agreement (DAA) in place. In most cases the expectation will be that anonymised data only is shared following the conclusion of the trial and the main analysis.

3. Roles and Responsibilities

The **Sponsor** of a research project as the Data Controller must ensure that an appropriate DSA or DAA is in place and provide guidance where requested. They may delegate responsibility to STU for any data sharing decisions relating to trial data.

The **Chief Investigator** (CI) as the research data custodian will be the recipient of the data request and will set in place an appropriate data sharing plan during trial conduct. If still available following completion of the trial, the CI will have the final say on the sharing of any data.

A **STU Data Sharing Committee** (DSC) consists of the STU Executive Team, STU IT Officer, a sponsor representative, the CI and other members of the Trial Steering Committee and Trial Management Group (where available). Data Sharing Request (DSR) forms received will be logged and Data Sharing Committee (DSC) convened. Where required, they will work in conjunction with the Sponsor, their Data Protection Officer (DPO) and research team committees (where still available) to review all data requests received, conclude a DSA/DAA and oversee the preparation and release of an appropriate data pack.

A **Data Manager** is responsible for reviewing and providing a dataset for the statistician to complete final cleaning.

A **Statistician** or delegate is responsible for preparation of data packs, dissemination to data requesters with approved data sharing agreements in place and liaising with data requesters as needed.

External use of SOP: this SOP and Associated Documents (AD) may be used for research projects not adopted by STU where Swansea University (SU) staff and associated NHS organisations require guidance. In such instances, oversight responsibility for any associated tasks will not be the responsibility of STU.

4. Procedure

Data sharing must be conducted in accordance with STU-SOP-TM-006 Data Protection.

4.1 Designing a Data Sharing Plan (DSP)

The DSP will provide project specific information on the management of research data for sharing. This will be in alignment with the research project protocol, the Patient Information Sheet (PIS) and participant consent given. Where a funding body requires a DSP their policy will be referenced in the DSP. The DSP can be drafted as a stand-alone document or incorporated into the Data Management Plan.

A DSP should detail:

- A description of all available datasets
- Data security and information governance
- Timescales for data release

4.2 Data Sharing Request Process

All data sharing requests should be made via STU@swansea.ac.uk by submitting a completed Data Sharing Request (DSR) form (STU-AD-FRM-038). When the DSR is received by STU, the request will be considered by the STU DSC. All initial enquiries for data sharing will receive a response within 5 working days. There will be an aim of releasing data within 28 working days, dependent on completion of a DSA.

Data sharing requests will not normally be approved before a trial has been concluded and the final report published.

Where the CI of the trial is available they will make the final decision on sharing data.

4.3 Review of Data Sharing Requests

DSRs are reviewed by the STU DSC. The DSR will be reviewed and any queries or issues raised with the data requester.

Consideration will be given to:

(i) the trial Data Sharing Plan (legacy projects may not have this in detail).
(ii) when the data will be available to share.
(iii) who has responsibility for authorising release of the data (e.g. Sponsor will have the authority to share the data but may have delegated to the CI/STU).
(iv) restrictions on sharing the data (e.g. consent restrictions on sharing, issues with data collection or coding that may impose restrictions).
(v) validity of the research proposal.
(vi) suitability of the person/organisation requesting the data (competence/qualification e.g. supervision if PhD student).
(vii) risks of sharing the data if not anonymised (e.g. participant identification, integrity of the trial) and how risks can be mitigated.
(viii) if data is removed from Swansea University, the data and security policies of the recipient organisation.
(ix) costs to STU, if any, which must be communicated to the requester with detail about what the cost covers and what will be recovered from the requester.
(x) Any additional criteria deemed relevant by the DSC.

4.4 Costs

STU will usually not charge for providing data packs to data requesters. However, for instances where the request requires additional coding or significant amount of work then the requester will be expected to pay reasonable costs for this work.

In such instances an invoice will be included as part of the DSC decision.

4.5 Decision

The decision to approve, or not, the data sharing request will be documented in the DSC meeting minutes. The committee can choose to:

- Approve
- Approve subject to additional conditions
- Refuse with an explanation provided

The decision will be communicated in writing to the requester within 10 working days of the meeting.

Should the requester wish to appeal the decision, appeals must be made within 28 days of the date of the decision communication, with all additional information included in the appeal. In the case of a dispute about the decision, the appeal will be escalated to the Sponsor of the original research project.

Requests to further share data will be subject to a separate DSA or DAA.

4.6 Data Access Agreement (DAA)

A DAA may be used for the release of anonymised data where a Data Requester will access research data which will remain on the STU/SU secure servers. Completion of a data access agreement removes the need for institutional sign off and data transfer requirements but will still require the completion of a DSR and agreement of the CI or DSC as appropriate.

The agreement must include considerations for safeguarding data, rules for future publication and academic credit, restrictions on subsequent data sharing, access to outputs and who should sign the agreement.

4.7 Data Pack

A Data Pack will only be prepared and released when a DAA or DSA is finalised and signed. The guiding principle is that the minimum amount of data in the least identifiable format should be shared.

The data pack should contain data and supporting documentation as appropriate:

(i) Dataset(s) (process for anonymisation is shown in Appendix 2)
(ii) Data dictionary
(iii) Additional metadata e.g. annotated Case Report Forms
(iv) Reference to any data standards used
(v) Relevant version(s) of the protocol
(vi) Extracts of the relevant Statistical Analysis Plan
(vii) Ways in which shared data differs from the published data if relevant (e.g. withdrawals)

The data pack will be generated by the trial statistician wherever possible with support from other STU staff as required. A quality check of the dataset should be completed by a member of STU not involved in data pack production to ensure that data can be understood and complies with the DSR.

4.8 Provide Access/Transfer of Data

All use, storage and handling of data must comply with data protection legislation and the DSA. It is the responsibility of the data requestor to ensure that adequate security levels are in place for data storage and processing.

The DSR specifies how the data requester wish to receive the data. Options for data transfer include:

- Secure electronic transfer by STU - Passwords to unencrypt the data should be provided to the recipient separately from the data e.g. by telephone.
- Post.
- Physical transfer by STU - Identifiable data should never be sent through a regular postal service or using unencrypted data transfer systems.

The STU Data Transfer Form (STU-AD-FRM-018) must be completed when a data pack is available for sharing. This must be logged in the STU Data Transfer log by the STU IT Officer or delegate responsible for the release of the data.

All methods used for data transfer should be appropriately tested to ensure compatibility with the transfer process.

4.8.1 By post

All personal identifiable data sent or received by post, email or on electronic removable media should use only approved methods e.g. encrypted memory cards, by registered post in tamper proof envelopes. All such deliveries should be logged as received along with the date of receipt and the name of the receiver.

Documents containing sensitive personal data or audio/video recordings of consultations or interviews should be labelled with only the unique study identifier and sent by registered post or courier. The DSP if available or the research project protocol will describe which postal methods should be used and advice sought from the CI/STU Manager or relevant DPO where necessary.

4.8.2 Electronic data transfer

Secure File Transfers which include encryption must be used. This may be done using STU systems or an external provider. The DSP if available or the research project protocol will describe which methods should be followed and advice sought from the CI/STU Manager or relevant DPO where necessary.

4.8.3 Outside of the European Economic Area (EEA)

Personal data shall not be transferred to a country or territory outside the EEA unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data. Consent must have been obtained from the data subject for the data to be transferred. A DSA must be in place between the research sponsor and the vendor / third party(ies) involved before any data transfer.

4.9 Transparency

To aid in the discoverability of STU datasets and the STU data sharing process the STU website will be used to:

- make the data sharing procedure available with direct links available to the Data Sharing Request Form (STU-AD-FRM-038)
- publish a summary of previous requests for data and decisions made
- publish a list of available datasets including an assigned Digital Object Identifier (DOI)

In addition, an internal log will be kept of all data request recipients.

4.10 Expectations on Data Requesters

Following an approved DSR and dissemination of data, data requesters are required to abide by the conditions of the DSA. In any reports or publications, acknowledge the contribution of the original research team and cite appropriate literature from the original trial in accordance with the DSA and academic standards.

5. References

- Good Practice Principles for Sharing Individual Participant Data from Publicly Funded Clinical Trials. Tudur Smith C, Hopkins C, Sydes M, Woolfall K, Clarke M, Murray G, Williamson P. April 2015.
- Health Research Authority website (HRA) - http://www.hra.nhs.uk/
- Medicine and Healthcare products Regulatory Agency website (MHRA) - https://www.gov.uk/government/organisations/medicines-and-healthcare-productsregulatory-agency/services-information
- UK policy framework for health and social care research (2017) - https://www.hra.nhs.uk/planning-and-improving-research/policies-standardslegislation/uk-policy-framework-health-social-care-research/
- UK Medicine for Human Use (Clinical Trials) Regulations 2004 - http://www.legislation.gov.uk/uksi/2004/1031/contents/made

It is assumed that by referencing the principal regulations that all subsequent amendments are included in this citation.

6. Associated Documents

Number	Title	Location
STU-AD-FRM-038	Data Sharing Request Form	Q-Pulse
STU-AD-FRM-018	Data Transfer Form	Q-Pulse
STU-AD-TMP-050	Data Transfer Log	Q-Pulse

7. Abbreviations

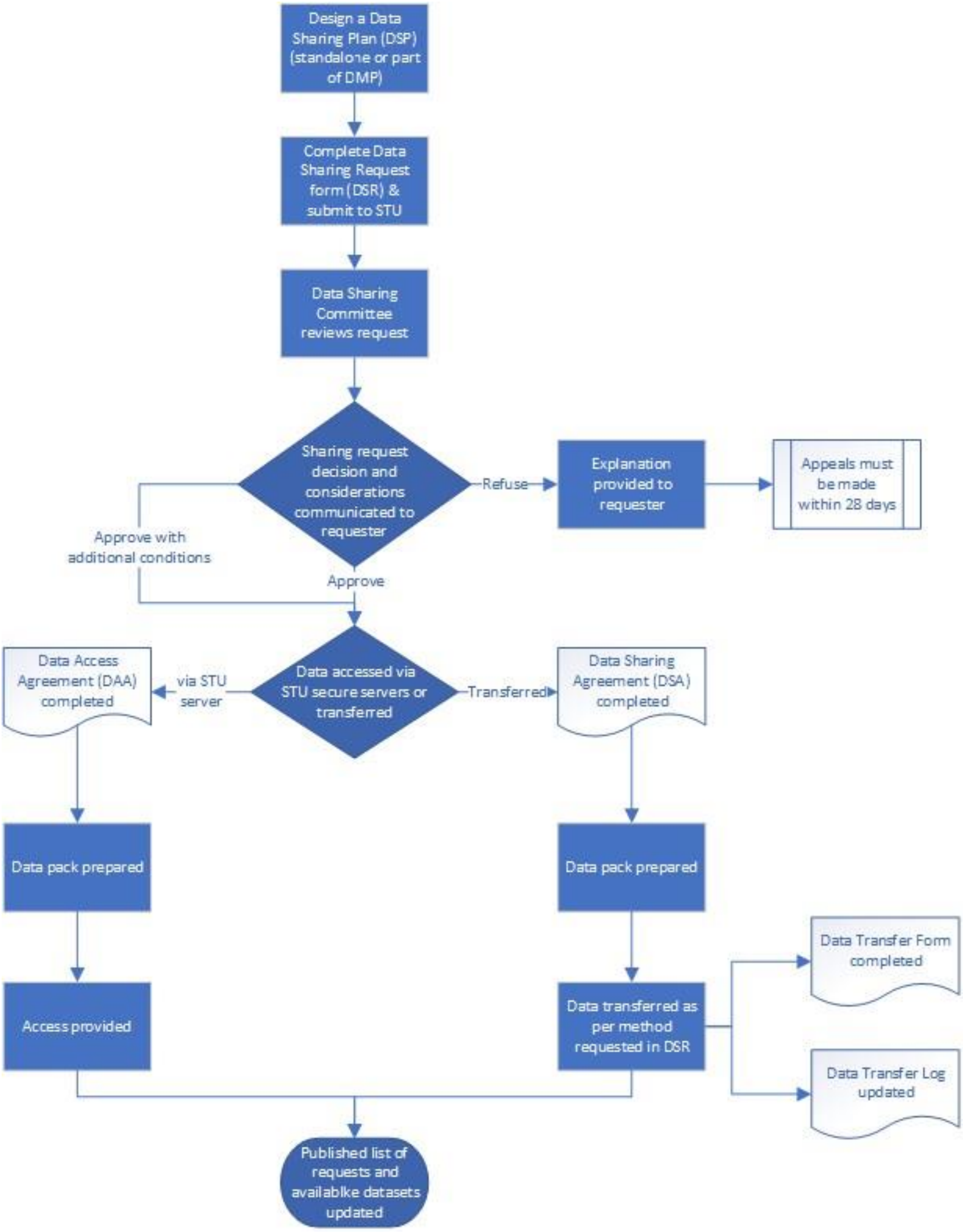
List of Abbreviations	
CI	Chief Investigator
DAA	Data Access Agreement
DC	Data Controller
DP	Data Processor
DPA	Data Protection Act
DPO	Data Protection Officer
DSA	Data Sharing Agreement
DSC	Data Sharing Committee
DSP	Data Sharing Plan
DSR	Data Sharing Request
EEA	European Economic Area
FoI	Freedom of Information
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
ICF	Informed Consent Form
ICO	Information Commissioners Office
ISF	Investigator Site File
PI	Principal Investigator
RGF	Research Governance Framework
SOP	Standard Operating Procedure
STU	Swansea Trials Unit
SU	Swansea University

8. Appendices

Appendix 1: Document History

Version No:	2	**Effective Date:**	26-Mar-2024
Description of changes:	Update to procedure and roles and responsibilities Updated to SOP Template v5		

Appendix 2: Data Sharing Request Flow Chart



Design a Data Sharing Plan (DSP) (standalone or part of DMP)

Complete Data Sharing Request form (DSR) & submit to STU

Data Sharing Committee reviews request

Sharing request decision and considerations communicated to requester

Refuse

Explanation provided to requester

Appeals must be made within 28 days

Approve with additional conditions

Approve

Data accessed via STU secure servers or transferred

via STU server

Data Access Agreement (DAA) completed

Transferred

Data Sharing Agreement (DSA) completed

Data pack prepared

Data pack prepared

Access provided

Data transferred as per method requested in DSR

Data Transfer Form completed

Data Transfer Log updated

Published list of requests and availablke datasets updated

Appendix 2: Anonymisation process

(Taken from the UKCRC Data Sharing SOP Guide v1.1)

ID Variables

- Recode unique patient identifier (ID) variables
- ID variables should be recoded consistently across datasets (and extension studies if relevant) to allow linkage
- Identify and recode other unique identifiers e.g. centre identifier variable

Convert Dates

- All dates (e.g. randomisation, clinic visit, date of adverse event, date of death) should be converted. Example methods:
  - Offset dates for each individual by subtracting a small number from each date for that individual. The small number may be generated randomly or calculated (e.g. date of first visit for the individual = screening date - anchor date for the trial = initiation date)
  - Convert all dates to project days since randomisation
  - Convert date of birth to age at randomisation
- In the case of partial dates, converted dates should remain partial

Personally Identifiable Information

- Remove or recode personally identifiable information (PII) and sensitive data. For example:
  - Remove patient name, address, initials
  - Recode place of birth, convert date of birth to age at randomisation
  - Consider PII of third parties e.g. remove investigator name

Extreme Values

- Investigate extreme values / rare characteristics
- Tabulate categorical variables and consider small cells, consider the minimum and maximum value of continuous variables and consider the risk of re identification
- If required, consider re categorising a variable, e.g. raise country to continent or use age categories rather than exact age, but consider the utility of transformed data

Text Variables

- Complete or partial removal of free text
- Consider whether free text variables are of clinical utility and whether the identifiable part of the text could be removed rather than complete redaction, e.g. ‘Dr X recommended a lower dose of the drug’

Quality Control

- Quality control checks performed by an independent person
- Review PII/ sensitive data, low frequencies and free text judgements
- Determine if removal of PII and sensitive data successful and if further removal required
- Re run basic analyses from the original project and compare results
- Check the accuracy and completeness of the document summarising anonymisation process