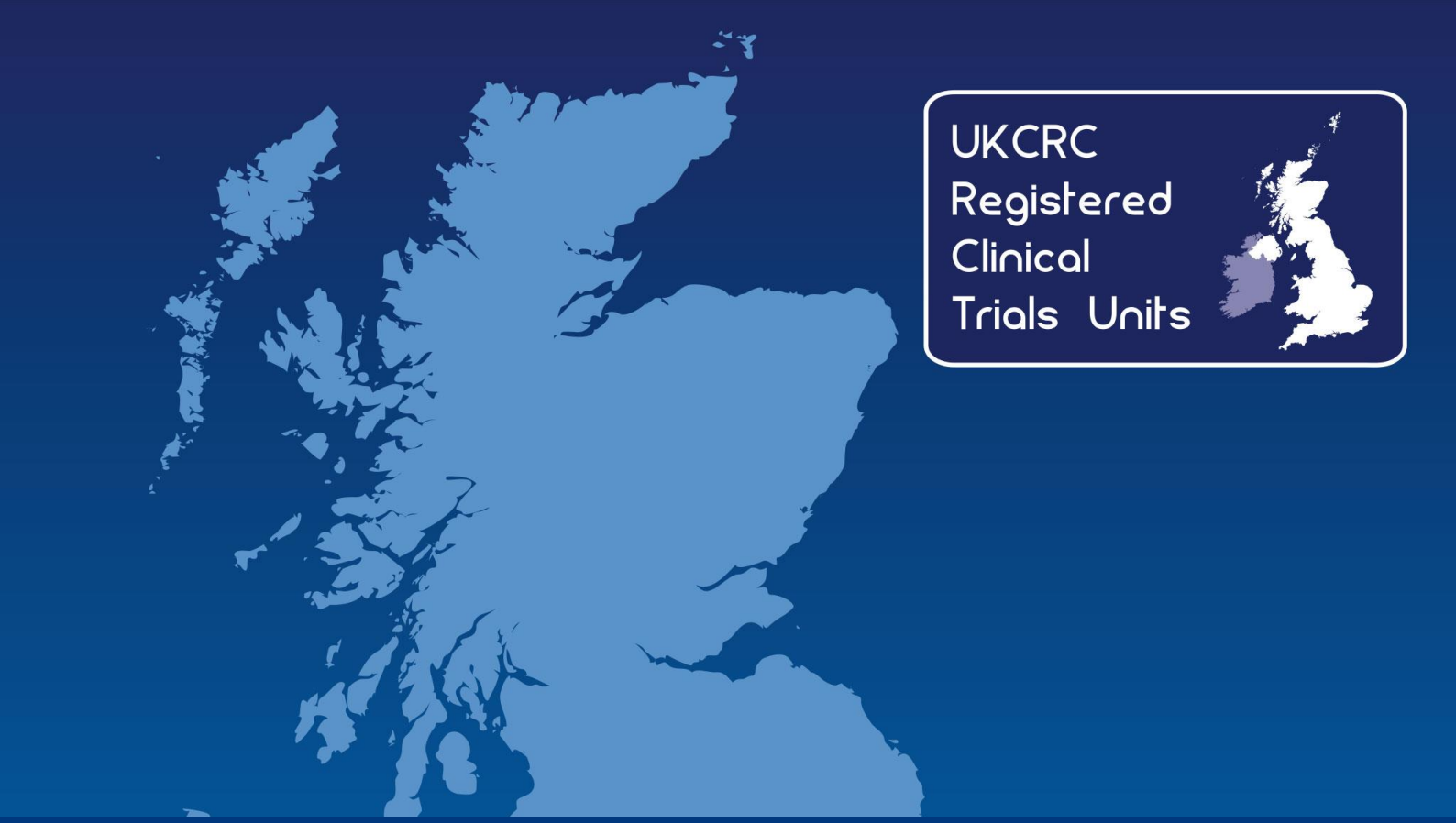

UKCRC Registered Clinical Trials Units

Considerations for a Participant Data Sharing SOP

UK Clinical Research Collaboration

igniting our potential

Contents

Considerations for a participant data sharing SOP

Scope and Background

This document has been compiled by the UKCRC Registered CTU Network Patient Data Sharing Task and Finish Group. The aim is to outline key items for consideration by a data owner (Sponsor or Clinical Trial Unit with delegated responsibility for data sharing) developing or refreshing a unit- or trial-level Standard Operating Procedure (SOP) for the process of sharing individual participant data (IPD) from clinical trials. The document covers the procedure from the point where a researcher (data requester) makes a request for data, to the point where the trial dataset is released. The primary focus is on sharing data after trial closure and for a secondary purpose but the principles may also be relevant to other scenarios. The considerations build upon previous recommendations for sharing publicly funded trial data (1) and should follow the FAIR principles as far as possible (2). Requests for data under the Freedom of Information Act 2000 are outside of the scope but are mentioned briefly.

NOTE: Links are correct at time of publication

Legal Considerations

Common Law Duty of Confidentiality

The Common Law Duty of Confidentiality applies to information about an identifiable individual who is alive or deceased, which is not in the public domain, which has a degree of sensitivity and was given the expectation it will be kept confidential. This includes medical and healthcare information.

In order to lawfully access confidential personally identifiable information such as that contained in health records for research activities, the access must be in line with the reasonable expectations of the individual about whom the information relates.

- For clinical trials and associated research projects, the default access route is reasonable expectations that is typically evidenced by written consent. Information will have been provided via routes such as the Participant Information Sheet, newsletters, and websites during the trial and this information can be used to evidence the reasonable expectations of participants of the research project.
- However, where it is not practical to obtain consent to access data held under a duty of confidentiality, there is an alternative route under the lawful provisions of

the NHS Act 2006 Section 251 that enables the Duty of Confidentiality to be set aside under limited and controlled circumstances. Such authorisation can only be provided by the Health Research Authority Confidentiality Advisory Committee (CAG).

The duty of confidentiality, once extended, continues after an individual has died, therefore confidentiality must be respected in line with the reasonable expectations or CAG Approval for as long as the information is held.

In most cases, data sharing requests can, and therefore should, be satisfied using an anonymised dataset. However, where a data sharing request requires an extension to the duty of confidentiality (i.e. the release of identifiable sensitive data), the custodian of the dataset (this could be the CTU host organisation or another contracting partner e.g. NHS Digital) must be satisfied that the request is in line with the reasonable expectations of the participant.

For more information, refer to:

- Health Research Authority Best Practice – Informing Participants and Seeking Consent https://www.hra.nhs.uk/planning-and-improving-research/best-practice/informing-participants-and-seeking-consent/
- Confidentiality Advisory Group - https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/confidentiality-advisory-group/
- Health and Social Care Information Centre Code of Practice on Confidential Information https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/codes-of-practice-for-handling-information-in-health-and-care
- Information Governance Alliance GDPR guidance on consent https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-detailed-guidance/legal-basis-processing-data/

2018 Data Protection Act

The 2018 Data Protection Act (3) applies to the processing of personal data, including special categories of personal data about people who are alive. It places certain obligations onto those organisations who are processing personal data (Data Processors) as well as those organisations who are determining why and how personal data is used (Data Controllers).

- **Personal Data –** means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
- **Special Categories of Personal Data –** personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation

Legal basis

To process special categories of personal data, a legal basis for the processing must be determined by the Data Controller. Typically, for research such as clinical trials, this will be Article 6e Task in the public interest and Article 9j (see reference 3), Archiving in the public interest, scientific or historical research or statistical purposes. The 2018 Data Protection Act requires that data processed using this legal basis must:

- Be in the public interest
- Be necessary for archiving purposes, scientific or historical research purposes or statistical purposes
- Not be likely to cause substantial damage or distress to the data subject
- Where the research involves making decisions which affect individuals such as in the case of interventional research like clinical trials the research must be approved by an authorised body such as an ethics committee
- Be protected with appropriate technical and organisational safeguards that ensure respect for the principle of data minimisation and exemptions to data subjects' rights are only exercised where this will render the research impossible or seriously impair it.
- Be in line with the transparency information provided to the data subjects'.

Rights and Exemptions for Research

Data subjects have certain rights under the 2018 Data Protection Act. However, there are some conditional exemptions to these for research where this would render the research impossible or seriously impair it. More information can be found here:

- Health Research Authority GDPR Technical Guidance: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-detailed-guidance/data-subject-rights-and-research-exemptions/

Sharing Data

In most cases data sharing requests can, and therefore should, be satisfied using an anonymised dataset. The assessment of anonymity should be considered from the perspective of the recipient, taking into account the measures taken to pseudonymise the data and other controls to prevent re-identification of the data, such as organisation policies, IT security and legally binding contracts.

Where Personal Data is to be shared, the Data Controller must satisfy itself:

- That the sharing would reasonably have been expected, based on the information given to the participants of the research project in line with the 2018 Data Protection Act transparency requirements.
- That the reason for sharing is compatible with the lawful basis for processing and will not be processed in a manner that is incompatible with those purposes. Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1), is not considered to be incompatible with the initial purpose but must take into account the transparency information given to participants.
- That any data shared is adequate, relevant and limited to what is necessary in relation to the purposes for which they will be further processed.
- That the data will be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.
- The Data Controller should keep appropriate records of all decisions made in relation to data sharing requests and associated activities to demonstrate compliance with the 2018 Data Protection Act.

Freedom of Information Act 2000

Requests for access to individual participant data obtained in the course of, or derived from, a research project is exempt where the project is continuing with a view to publication and the disclosure of the information would be likely to prejudice the project.

The Duty of Confidentiality and 2018 Data Protection Act would still need to be applied to any request made under a Freedom of Information request.

Where to get further information or advice:

- Host Institution Data Protection Officer or Freedom of Information Office
- Information Commissioners Office website

- MRC Regulatory Support Centre website

Contracting

Other contractual definitions and controls must be assessed and complied with prior to sharing data for a specific request. Data could be subject to intellectual property or ownership restrictions such as those applied by the original trial funder or data provider.

Contributors

Gill Booth – Leeds Clinical Trials Research Unit

Stuart Douglas – Public Contributor

Kate Fenton - NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC)

Carrol Gamble 0000-0002-3021-1955 – Liverpool Clinical Trials Centre

Julie Hepburn – Public Contributor

Nicola Howe 0000-0001-9446-8314 – Newcastle Clinical Trials Unit

Sharon Kean – Glasgow Clinical Trials Unit

Sally Kerry – Pragmatic Clinical Trials Unit

Steff Lewis – Edinburgh Clinical Trials Unit

Matthew Sydes 0000-0002-9323-1371 – MRC CTU at UCL

Catrin Tudur Smith (Chair) 0000-0003-3051-1445 – Liverpool Clinical Trials Centre

Appendix 1 - SOP Content Areas

SECTION OF SOP	CONTENT
1. Roles and Responsibility	Outline clear responsibilities and identify specific roles for CTU staff members involved in different activities. For example - Who will administer the initial request and liaise with data requester - Who will coordinate the review of the request - Who will identify and locate the dataset and relevant information - Who will liaise with data recipient's legal representatives to finalise the data use agreement - Who will prepare the data pack if not already prepared (trial team in most situations) Recommendation: A central single point of contact at the CTU should be used to administer and log requests and formal enquiries
2. Request process	Specify the process for data requesters to follow when making a request for data from the CTU. For example, - Methods for requesting data (e.g. through an online form, email contact, HDR Innovation Gateway) - The information required from the data requester (e.g. research proposal – see example) - What happens when the request is received (diagram) - Suggested timescale for processing the request

SECTION OF SOP	CONTENT
3. Review of request	Outline the process for reviewing data requests. For example, - Who will conduct the review (e.g. internal committee; committee with some independent members; independent committee). Membership of the review committee should be described and consideration given to PPI contribution where relevant. Efficiency gains may be achieved by having a single committee. - What are the criteria for review, to include consideration of (i) If / when is data available to share? (ii) Is it the CTU responsibility to share? – trial sponsor has authority to share but this may be delegated to CTU (iii) What are the restrictions on sharing the data – (e.g. does the original PIS and consent form impose any restriction on sharing? Are there issues with the way data have been collected or coded that may impose restrictions on sharing? (iv) Scientific validity of the research proposal (v) Suitability of the person requesting the data (competence / qualifications e.g. support if a PhD student) (vi) What are the risks involved in sharing the data (e.g. Risks of participant identification; risks to the integrity of the trial); and methods to mitigate risks.
4. Decision	Include details of a transparent decision making process, including - Potential decision outcomes - Information to the requester to support any refusals - Appeals process - Escalation process in case of a dispute - Decision making process may need to consider costs which should be made clear to the requester and be transparent about what this cost covers

SECTION OF SOP	CONTENT
5. Prepare data pack	- Wherever possible, the minimum amount of data in the least identifiable format should be shared - Data Packs to include: ~ Anonymised dataset(s) (see Anonymisation Flow Chart) ~ Data dictionary ~ Relevant version(s) of the protocol ~ Relevant extracts of the relevant Stats Analysis Plan ~ Ways in which shared data differs from the published data if relevant (e.g. withdrawals) ~ Blank case report forms (if relevant) - It is good practice to include a quality check of the dataset by an individual not involved in the data pack preparation to (i) check that data are anonymised, (ii) to check that data can be understood - Sharing complex datasets or those without adequate supporting information may require additional information to be provided to data recipients to minimise errors / misunderstandings
6. Data Use Agreement	If using data use agreement (recommended) outline the content and process (include template), to consider - Safeguarding data and re-identification - Rules relating to future publication - Academic credit - Monitoring of future outputs - Restrictions on subsequent sharing - Who should sign contract
7. Provide Access/ Transfer	Reflect the different mechanisms available for access/transfer: - Sponsor's data safe haven - Other external data safe haven - Secure electronic transfer by CTU - Physical transfer by CTU

SECTION OF SOP	CONTENT
8. Transparency	Consider approaches to increase discoverability of datasets and transparency of the data sharing process e.g. - Assign a Digital Object Identifier (DOI) on the dataset - Publish the data sharing process and a list of available datasets (e.g. on the CTU website) - Publish a summary of previous requests and decisions for datasets - Maintain a log of data requests and progress of data recipients

1. Tudur Smith C, Hopkins C, Sydes MR, Woolfall K, Clarke M, Murray G, et al. How should individual participant data (IPD) from publicly funded clinical trials be shared? BMC medicine. 2015;13:298.
2. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 2016;3(1):160018.
3. Data Protection Act 2018. Available at http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted (Accessed: 23 December 2020).

Appendix 2 - Preparing Anonymised Individual Participant Data

Introduction

About this guidance

The following outlines the key steps when anonymising individual participant data from a clinical trial in preparation for sharing with external researchers within a 'controlled access system' (i.e. data access requires approval and compliance with a formal data use agreement). The process is likely to require clinical input.

Using this guidance

Before the anonymisation process:

- **Identify all relevant study datasets and documentation.** Protocol, annotated case report forms (CRFs), visit schedule, data dictionary, statistical analysis plan etc.
- **Check original trial consent form for any data sharing restrictions.** Consider whether the dataset may contain data from non-trial participants e.g. a partner in a HIV trial. Note that it is not necessary to obtain patient consent to anonymise the data.
- **Prepare document to summarise anonymisation process to accompany the dataset.** This should include a description of the anonymisation methods and summary of variables that have been removed or redacted. Note that the unique patient identifier code and date of randomisation can provide valuable information about the sequence and pattern of randomisation that may not be possible to uncover in the anonymised dataset. A summary of flow of trial participants highlighting any randomisation errors could be provided to supplement the dataset.

The key steps of the anonymisation process

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 as (date of first visit for the individual e.g. screening date – anchor date for the trial e.g. initiation date) (see appendix 1, PhUSE de-identification standard, ref 4)
  - Convert all dates to study days since randomisation
  - Convert date of birth to age at randomisation
  - In the case of partial dates, converted dates should remain partial (see appendix 1, PhUSE de-identification standard, ref 4)

Personally Identifiable Information

- Remove or recode personally identifiable information (PII) and sensitive data. Form example:
  - Remove patient name, address, initials
  - Recode place of birth, convert date of birth by 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'

Appendix 2 Resources

1. Tudur Smith C, Hopkins C, Sydes MR, Woolfall K, Clarke M, Murray G, Williamson P. How should individual participant data (IPD) from publicly funded clinical trials be shared? BMC Med. 2015 Dec 17;13:298.
2. El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. BMJ. 2015 Mar 20;350:h1139.
3. Hrynaszkiewicz I, Norton ML, Vickers AJ, Altman DG. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. BMJ. 2010 Jan 28;340:c181.
4. PhUSE De-Identification Working Group. De-Identification Standard for CDISC SDTM 3.2 version: 1.01. 20MAY2015 http://www.phuse.eu/data-transparency-access
5. Information Commissioner's Ofice. Anonymisation: managing data protection risk code of practice. https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf