NHS
The Leeds Teaching Hospitals
NHS Trust

Research Data Access Request - Guidance

Author: Research Data and Informatics Team

Date: December 2022

Version: 2.0

Policy Title	Research Data Access Requests - Guidance
Version:	2.0
Data Access Request form version:	10.0
Approved by:	John Corkett
Date of approval:	2022-12-08

TABLE OF CONTENTS

Research Data and Informatics Team

The Research Data and Informatics Team (R-DIT) provides informatics support within the Research and Innovation department. The team also delivers a Data Access Request Service to provide datasets for research projects and approval for using data for secondary purposes where the data are not extracted by R-DIT. This guidance document describes the process of the Data Access Request Service.

1 Research Data Access Request

To request data for research purposes, a Data Access Request Form is completed and reviewed by the Data Access Committee (DAC) for approval (Figure 1). The Data Access Committee includes representation from the Caldicott Guardian, the Trust's Data Protection Officer, the Senior Information Risk Owner and Research and Innovation. Research Governance also review forms that come under the umbrella of an HRA approved study.

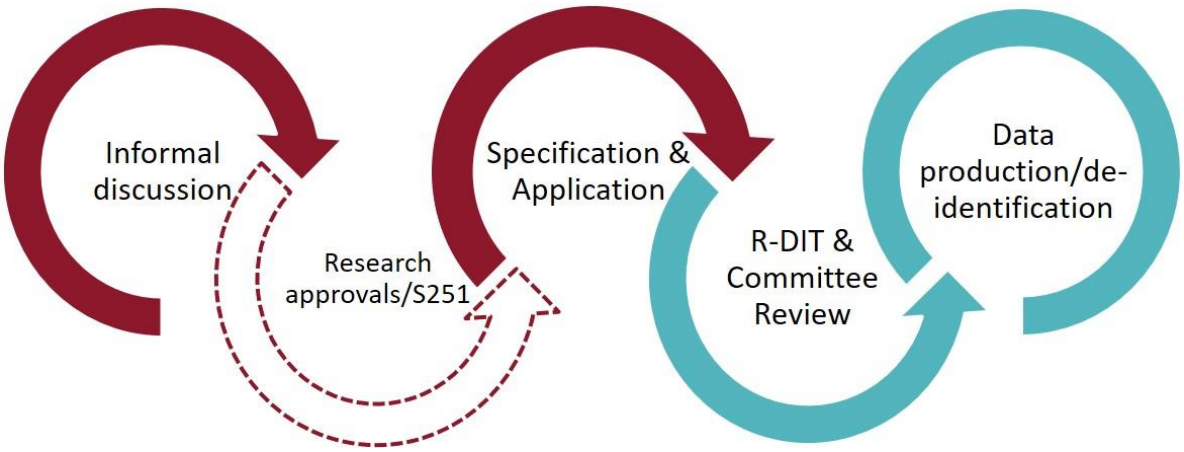

Figure 1: Research Data Access Request process.

1. **Informal discussion**
   You must contact R-DIT prior to completing a Research Data Access Request Form (DARF). We will arrange an initial meeting to discuss your requirements, including the de-identification process and give guidance on how to proceed with your funding. Other topics to discuss will include what approvals you have, are you linking to a patient cohort, are all the data points known etc.?

   You can email the team on leedsth-tr.researchdata@nhs.net. If you require further guidance or have any questions regarding research data, please feel free to get in touch at the same email address.

   Prior to any release of data, the requester and any persons who will access the data must declare they have attended Mandatory Information Governance Training or the Research Academy course: "How to Use Real World Data from LTHT for Research". Equivalent training in another organisation is acceptable. Please check our website for upcoming dates.

   Research approvals and Section 251
   To obtain personal data (What is personal data? | ICO) you will need either:
   - Patient consent
   - Approval from the Confidentiality Advisory Group for Section 251 approval

   If you do not have either of these approvals, you can only access data that is either anonymised or de-identified.

2. **Application**

There are four types of application:

- **New**
  The first time an application is received for a specific project
- **Extension**
  A request to retain the data that is held for longer than originally stated
- **Renewal**
  A request for the exact same data points as provided for the project previously but for a different time period
- **Amendment**
  A change to what was previously requested for a specific project in terms of storage location, who will access the data and any changes to the dataset.

Please refer to 1.1 Research Data Access Request Form when completing the form. You will work with the R-DIT analyst to put together the application. You will need to accept the recommendations that the analyst gives you and if you feel these are not appropriate you will need to justify these points in the application. This helps to get the application approved first time wherever possible.

Before an application for using the data can go to the DAC a data specification must be completed. You will have the responsibility of working with the analyst in the R-DIT to put the specification together. In addition, the R-DIT must have access to the data sources required for the project.

It is expected that you will be responsive to requests for information during these activities. It is your project, and you must drive it. It is not the R-DIT's job to chase you for requested information. If there is no response in a reasonable amount of time, the R-DIT reserve the right to cancel your request. It will be assumed that the project is not important to you. If there are valid reasons for delays, you must let us know and we will be happy to put the request on hold for a period.

You must ensure that all necessary research approvals are in place or in the process of being approved if your request involved HRA or CAG. A Data Protection Impact Assessment (DPIA) has to be completed before submitting your application. In the case of de-identified/anonymous data, this will no doubt require completion of the screening questions only. Once completed, submit the form via email to R-DIT on leedsth-tr.researchdata@nhs.net.

3. **Committee Review**

The form will be reviewed by the Data Access Committee. Once a decision has been reached, R-DIT will provide you with feedback of one of the following:

- Approved
- Approved with caveats
- Not approved and reasons why

If the request is rejected and you wish to continue with your project, you will need to reapply with a new Research Data Access Request submission for the next Data Access Committee meeting. Please make sure all points of feedback from the previously

rejected request are addressed in the new application. If the application is approved, you will receive a letter of approval via email.

Data Extraction/De-identification

Before work commences on your project, how you pay for the service will be confirmed, usually in the form of a Purchase Order number.

The goal is to provide data within 90 days of approval. The R-DIT will put work on hold after approval if assurance of payment is not received or a required cohort has not been received. When these are received the clock will start. The 90 days applies to the first dissemination of data only. You will need to agree in advance what this will include as part of the specification process. There will be no changes to what has been agreed in the DARF, the specification and any caveats applied by the DAC.

4. **Data Destruction**
   Data must be destroyed, and a certificate of destruction completed when the date of your project end date has been reached. If you wish to retain the data for longer you must apply for an extension.

1.1 Research Data Access Request Form

Details to consider whilst completing the form:

Has funding been secured for the project? **If YES, then what is the funding source?**	R-DIT works on a cost recovery basis and there is a cost associated with every application. If the application doesn't require much work and is internal to LTHT you may not be charged the full amount.
Project Title	Title of Research Project
Description of Project (250 words)	Description of Project: • What is the aim of the project? • Why is there a need for this project? • What outcomes are being measured? • How will these outcomes be measured? • What are the research questions?
Lay Summary (250 words)	A description of the project for non-clinical persons. This is required for a better understanding of the project for non-clinical members of the Data Access Committee, R-DIT and Information Governance, as well as members for members of the public so they can understand the use of patient data for research.
Patient benefits (200 words)	List the key patient benefits this project will create for current and future patients.

Is aggregated data required or record level data?	State whether the request involves aggregated numbers only (i.e. a table of numbers) or record level data (i.e. a dataset comprising of rows of data representing an individual record, such as an individual patient).
Datasets required for project	Unless you have patient consent or S251 approval the data you will be approved to receive will be either anonymised or de-identified (see section '2 Considerations when requesting for Research Data' for definitions). Provide a detailed list of data variables/fields, filters and systems (if known) required for the final dataset of your project. Dataset • If known, state which system(s) the data is required from e.g. PPM+, PAS etc. • If not known then describe the data sets that you want to receive, e.g. admissions, outpatients, chemotherapy, diagnoses, procedures, etc. Fields • Consider if data can be supressed or generalised e.g. values less than 5 replaced with '-'. • Variables such as age and exact dates are considered identifiable and may be rejected by the committee without good justification. If you need to include these variables in your request, consider how these variables can be grouped, e.g. 5-year age group, quarter of admission. Filters • What period the data is required e.g. from 1st January – 31st December 2019 inclusive. • If you require any specific diagnosis or procedures, please list the ICD10 or OPCS codes. These codes should be confirmed by the Clinical Coding team (leedsth-tr.clinicalcodingadvice@nhs.net) to enable all necessary codes are captured. • For filters such as specific age ranges, please specify if age at diagnosis is required or current age. • Is there any sequence to the order of events, e.g. an emergency admission before or after a cancer diagnosis? • Are there any exclusions that need to apply, e.g. sex, age, geographical area, emergency admissions only etc.? Derived cohorts R-DIT can create cohorts for the researchers based around particular diagnoses, procedures and so forth. Please provide the team with the necessary dataset and filter requirements as described above. Customer Cohort In some cases, applicants may send us a cohort for us to link to. In order to efficiently process this please supply the following information within this section of the form: • Approx. size of cohort e.g. 100 patients • Point of contact for queries relating to the cohort • What identifiers are within the cohort • Format the data is in • Where the data is currently stored and who currently has access • How the details of the cohort will be transferred to R-DIT securely

	Please note that national data opt-outs (the option for patients to opt-out of their data being used for research or planning) will be applied on all cohorts unless the patients have consented.
If personal data, which includes pseudonymised data, is requested. Is your use of the data consented by the patient or do you have a S251 from CAG? **If you have a S251 approval, please provide a copy of the letter.**	Ensure that all necessary research approvals are in place or in the process of being approved. If you are seeking anonymised or de-identified data do not fill out this section. If you have patient consent and do not require R-DIT to carry out your data extraction then you don't have to complete a DARF. If data is leaving LTHT then a Caldicott Letter is required.
Estimated size of data set (i.e. number of patients/ records/ occurrences etc.)	State the approximate size of the cohort if known. Small numbers may be classed as identifiable.
Data Flow Diagram	Provide a narrative description of the data flows for which support is sought. This must include: • The data sources • Relevant organisations • Identifiers to be processed • Stages where information is pseudonymised Applicants should also provide a simple graphical data flow diagram that shows the flow of information, who is involved in processing, the stages at which it is pseudonymised/anonymised.
Where will datasets be stored for the duration of the project? How will data be transferred if necessary?	Once data is disseminated where will this data be stored and for how long? See section '2.3 Data Security'. Storage options include NHS OneDrive but not personal or university OneDrive for example. Storage outside the UK will require more justification. If the data is to be transferred, please provide the name of the medium. Transfer options include: • NHS.net to NHS.net email accounts for small datasets • Secure File Transfer Protocol (SFTP) for large datasets • Any other methods must be approved by IG
Who will have access to the datasets for the purposes of the project (please give name, job title, organisation)?	Please list all the contacts who will be using the data once disseminated. Include the name of the organisation with which they are primarily affiliated and any honorary contracts they may hold with LTHT.
Have all persons who will have access to the data undergone LTHT Mandatory IG Training, attended the Research Academy course: "How to Use Real World Data from LTHT for Research" or both?	Prior to any release of data, the requester and any persons who will access the data must declare they have attended Mandatory Information Governance Training (or local equivalent) or the Research Academy course: "How to Use Real World Data from LTHT for Research". Please check our website for upcoming dates.

Will the datasets be linked with any other datasets (e.g. primary care, national datasets)? **If yes, please give details**	Please provide any details of any other datasets which you hold or will be linked to the data required.
Will any data be transferred to a third party?	State who the third party will be receiving data and what data will be transferred. Provide any additional details of how the data will be transferred securely.
How will the results of the project be used?	Include details of any upcoming events and meetings where you will intend to share the results of the project, as well as intentions for publication.
If any commercial organisations are involved in the project, how will they benefit from it?	Include any products or services the organisation will use with the data obtained.

2 Considerations when requesting for Research Data

2.1 Lawful Basis

If a data request is requiring personally identifiable data (PID), then a lawful basis for processing is required. To approve a data application which requires personal data (this includes pseudonymised data under General Data Protection Regulation (GDPR) you must have a lawful basis.

For further information please visit the Information Commissioner's Office website.

2.2 Data Minimisation

The General Data Protection Regulations state: "Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (data minimisation)"
General Data Protection Regulation Article 5 (1)(c)

This means any data requested must be justified by its purpose within the application, the datasets requested should be limited to only what is required for the project.

Even if data is not considered personal data, at LTHT we still require consideration to be given to this aspect. The more data items there are in a de-identified dataset, the greater the risk of re-identification.

Considerations should be made to the following aspects:

- Years
  Number of years should be limited to only what is requested and must be justified within the application
- Datasets
  Only the required datasets, and if identifiable data is being requested can this be achieved without the need for identifiable information?
- Filters
  To minimise the data required apply filters to narrow the search for example demographics such as age or geography.
- Variables
  Only request the required variables for your project. Clear dates and sensitive identifiers will most likely not be approved by the committee.

2.3 NHS Research

If you have HRA approval for a project but want to use patient data without consent, then you must complete a Data Access Request Form. It is important to recognise that the HRA are not accountable for how Trust data is used in law. Even though the HRA indemnify the organisation (Indemnity cover for NHS staff delivering research - Health Research Authority (hra.nhs.uk)) for using data for research, the Trust still has a legal, moral and ethical responsibility to ensure our patients' data is used appropriately and securely. Insurance cover does not excuse carelessness and negligence.