

# DATA SHARING GUIDANCE FOR CRUK RESEARCHERS

---

## INITIATIVES AND REPOSITORIES TO SUPPORT CLINICAL RESEARCHERS WITH DATA MANAGEMENT AND SHARING

### CONTENTS

1. Introduction .....	1
2. Generalist data sharing initiatives and repositories .....	2
2.1. Initiatives.....	2
2.2. Repositories .....	2
3. Initiatives and repositories specific to data sharing in clinical research.....	3
3.1. Initiatives.....	3
4. Good practice documents for clinical research data sharing .....	4

## 1. INTRODUCTION

Clinical trial data sharing has challenges, such as ensuring patient confidentiality and appropriate re-use of data by third parties.

Initiatives such as [ClinicalStudyDataRequest.com](https://clinicalstudydatarequest.com) – whereby pharmaceutical trial sponsors provide anonymised individual patient data to researchers whose proposals are assessed by an Independent Review Panel - demonstrate that there are feasible systems for secure, controlled sharing of individual patient data from industry trials.

In the following pages, we highlight some key initiatives and repositories which may provide tools and guidance for data sharing in the academic clinical trial field. Some of these are not discipline specific and are described as “generalist” ([Section 2](#)); then [Section 3](#) and [Section 4](#) respectively outline initiatives and good practice guidelines supporting data sharing in the Clinical Research field. It is the responsibility of the investigators to ensure that any repositories/standards/tools they intend to use are appropriate for the nature of the research envisaged.

### CONTACT US

We would be grateful for any comments or suggestions to help improve this guidance. Please get in touch with [jamie.enoch@cancer.org.uk](mailto:jamie.enoch@cancer.org.uk) with any feedback.

## 2. GENERALIST DATA SHARING INITIATIVES AND REPOSITORIES

### 2.1. INITIATIVES

Listed below are generic initiatives supporting data sharing in science and health research generally, which may provide useful tools, resources and methods you can factor into your data sharing plan.



[DataCite](#) provides advice on creating digital object identifiers (DOIs) for datasets. It also offers a range of services including a [Metadata search](#), which allows researchers to locate individual datasets through access to the relevant metadata, and a [search tool](#) to discover appropriate research data repositories worldwide through [Re3Data](#).



The [Digital Curation Centre](#) provides expert advice and practical help to researchers to store, manage, protect and share digital research data. It maintains a range of resources including [How-to Guides and checklists](#), [case studies in research data management](#), and [training programmes](#) for researchers and data custodians in research data management and sharing.



The [Expert Advisory Group on Data Access](#) convenes leading researchers on issues of data access and sharing who advise the funders on technical and legal issues in data sharing. It has produced a number of reports which advise on good practice in data sharing policy and governance.



[GigaScience](#) is an open-access open-data journal for ‘big data studies’ from across the life/biomedical sciences, in collaboration with BioMed Central. In many ways it is half journal and half repository. It links a standard manuscript publication with an extensive database, the [GigaDB](#), which hosts associated data and provides data analysis tools and cloud-computing resources.



[Nature Scientific Data](#) publishes peer-reviewed articles known as data descriptors, which focus on helping others reuse data by describing the dataset with structured, machine readable information. The articles can be descriptions of datasets of any size, and can link to datasets underpinning published research or describe standalone datasets. The journal mandates the release of datasets accompanying manuscripts, and links to datasets hosted on third-party repositories.



The [Research Data Alliance](#) works to build the social and technical bridges to facilitate data sharing and re-use. Its constituent [working groups](#) aim to tackle fundamental issues in data sharing; for example, [one working group](#) is seeking to rationalise databases, standards and funder policies in the biomedical sciences by working with the BioSharing platform.

### 2.2. REPOSITORIES

While good practice in clinical trial data sharing generally favours a managed data access approach, these generalist repositories may be appropriate for the storage of de-identified or aggregated data.

REPOSITORY	TYPES OF DATA RESEARCHERS MAY SUBMIT
<a href="#">Dryad Digital Repository</a>	Data underlying scientific or medical publications
<a href="#">Figshare</a>	Various data types, including figures, datasets and images
<a href="#">GigaDB</a>	Data and tools for <i>GigaScience</i> and other articles
<a href="#">Zenodo</a>	Research outputs from all fields of science

### 3. INITIATIVES AND REPOSITORIES SPECIFIC TO DATA SHARING IN CLINICAL RESEARCH

#### 3.1. INITIATIVES

Examples of initiatives with relevance to academic clinical trial data sharing are shown below.



The [Clinical Data Interchange Standards Consortium](#) aims to develop and support global, platform-independent data standards that enable interoperability between clinical information systems. The free-to-use [standards](#) support the acquisition, exchange, submission and archive of clinical research data and metadata.



The [Data Sharing & Transparency Initiative at the Multi-Regional Clinical Trials Centre](#) of Brigham and Women's Hospital and Harvard is spearheading the development of a new platform to provide secure federated access to individual patient data from academic clinical trials. This builds on the recommendations of a seminal 2015 US Institute of Medicine report on [Sharing Clinical Trial Data: Maximising Benefits, Minimising Risks](#).



[Project Data Sphere](#) provides a platform to share, integrate and analyse patient-level, comparator arm, phase III cancer data (which providers are required to de-identify). Protocols, data descriptors and case report form templates are provided on the website to enable users to tap into the value of the data.



Registries, including [ClinicalTrials.gov](#), [ISRCTN](#) and the [EU Clinical Trials Register](#), play a key role in allowing trials and their datasets to be discovered. The [AllTrials campaign](#), which is supported by CRUK, is aiming to ensure that all trials are registered with a summary of the trial protocol before the first patient is recruited.



The [Yale Open Data Access Project](#), which advocates for the responsible sharing of clinical research data, open science and research transparency. YODA works with pharmaceutical companies such as Johnson & Johnson and Medtronic, providing them with technical expertise to share their clinical trial data securely with bona fide researchers. It also has a [useful list of literature](#) relevant to clinical trial data sharing.

#### 4. GOOD PRACTICE DOCUMENTS FOR CLINICAL RESEARCH DATA SHARING

Although there are currently no repositories specifically for storing/sharing academic clinical trial data, the following documents provide guidance on good practice in managed access approaches to data sharing:

- Hrynaszkiewicz, I., Khodiyar, V., Hufton, A. and Sansone, S-A. 2016. [\*Publishing descriptions of non-public clinical datasets: guidance for researchers, repositories, editors and funding organisations\*](#)
- Institute of Medicine. 2015: [\*Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk\*](#)
- Medical Research Council. 2012.: [\*Policy on sharing of research data from population and patient studies\*](#)
- Tudur Smith, C., Hopkins, C., Sydes, M., Woolfall, K., Clarke, M., Murray, G. and Williamson, P. 2015. [\*Good Practice Principles for Sharing Individual Participant Data from Publicly Funded Clinical Trials\*](#), which has been given support by the National Institute for Health Research
- Wellcome Trust/Technopolis. 2015: [\*Assessing the research potential of access to clinical trial data\*](#)