

Cancer Research UK Data Sharing Guidelines

Cancer Research UK's Stance on Data Sharing

Cancer Research UK is committed to ensuring that the data generated through its funding should be put to maximum use by the cancer research community and, whenever possible, is translated to deliver patient benefit. It is therefore the policy of Cancer Research UK that all data generated as a result of its funding be considered for sharing and made as widely and freely accessible as possible whilst safeguarding intellectual property, the privacy of patients and confidential data. Scientists applying for funding should familiarise themselves with Cancer Research UK's Data Sharing and Preservation Policy. Further information and guidance on data sharing, databases and data standards is available at www.cancerinformatics.org.uk.

Given the diverse nature of the research Cancer Research UK supports, the guidelines below do not prescribe precisely how and when investigators should share research data. Instead they should be used to ensure that the principles of the policy are adhered to.

Applicability

Cancer Research UK's Data Sharing and Preservation Policy is applicable to all candidates seeking funding from Cancer Research UK after 1st April 2009 and applies:

- To the sharing of final research data for research purposes.
- To basic research, clinical studies, surveys and other types of research supported by Cancer Research UK.
- Especially to unique data that cannot readily be replicated.
- To projects that transform or link pre-existing datasets.

The data from all activities in the preparation for and arising out of Phase I and II clinical trials which Cancer Research UK sponsors and which are initiated by its Drug Development Office, after approval by its New Agents Committee, is not automatically covered under this Data Sharing Policy. For clarity on the position, please contact the Drug Development Office on a trial by trial basis.

Data Management and Sharing Plan

All applicants seeking funding from Cancer Research UK will be required to submit a data sharing plan as part of their research grant proposal. If data sharing is not appropriate, applicants must include a clear explanation why. The data sharing plan will be reviewed as part of the funding decision. Funding Committees will assess the appropriateness and adequacy of the data sharing plan and provide specific feedback to applicants where necessary.

Cancer Research UK recognises that data sharing strategies will vary according to the type of data collected and thus does not specify the exact content and format of the data sharing plan. We recommend that data should be shared using established standards and existing resources where possible. The following should be considered when developing a data sharing plan:

- The volume, type, content and format of the final dataset
- The standards that will be utilised for data collection and management
- The metadata, documentation or other supporting material that should accompany the data for it to be interpreted correctly
- The method used to share data

- The timescale for public release of data
- The long-term preservation plan for the dataset
- Whether a data sharing agreement will be required
- Any reasons why there may be restrictions on data sharing, for example
 - Development arrangements through Cancer Research Technology including intellectual property protection and commercialisation
 - Proprietary Data - restrictions due to collaborations with for profit organisations
 - International policies governing the sharing of data collected outside of the UK
 - Confidentiality, ethical or consent issues that may arise with the use of data involving human subjects.

Funding committees will monitor investigators' progress in implementing their data management and sharing plan. However, Cancer Research UK understands that an investigator may need to adapt the method and timelines for sharing during the course of the study - for example when potential intellectual property arises unexpectedly.

Intellectual Property Rights and Proprietary Data

Data which might have the potential to be exploited commercially or otherwise to deliver patient benefit should be discussed with your technology transfer office and Cancer Research Technology prior to data sharing.

Cancer Research UK encourages the appropriate filing of patents and recognises that there may be a need to delay the release of data until patent applications have been filed. Whilst there may be a delay in the release of data due to the application process, appropriate intellectual property protection should not hinder data sharing and may be the best way of ensuring that patient (and public) benefit is delivered.

Any intellectual property issues or plans for commercialisation that may affect data sharing should be addressed in the data sharing plan. Cancer Research UK understands that unexpected intellectual property may arise during the course of the study and investigators may need to depart from their data sharing plan to protect intellectual property and for any other necessary steps to be taken.

Data sharing may also be affected when co-funding is provided by the private sector (e.g. by a pharmaceutical company) or host institution resulting in some restrictions on the disclosure of data. For example with clinical trials, the Trial Management Group and/or trial sponsor etc may impose restrictions on data access. Any restrictions should be outlined in the data sharing plan and applicants should explore ways data sharing requests can be considered by the body that owns the data.

Standards, Metadata and Documentation

For data sharing to be a success it is important that data are prepared in such a way that those using the dataset have a clear understanding of what the data mean so that they can be used appropriately. To enable this, applicants are encouraged to include with the dataset all the necessary information (metadata) describing the data and their format. This information should include such information as the methodology used to collect data, definitions of variables, units of measurement, any assumptions made, the format of the data, file type of the data etc. To support this researchers are strongly encouraged to utilise community standards to describe and structure data, (e.g. common terminology, minimum information guidelines and standard data exchange formats).

Methods for Data Sharing

The methods used to share data will be dependent on a number of factors such as the type, size, complexity and sensitivity of data. Data can be shared by any of the following methods.

1. Under the auspices of the Principal Investigator

Investigators sharing under their own auspices may securely send data to a requestor, or upload the data to their institutional website. Investigators should consider using a data-sharing agreement (see below) to impose appropriate limitations on the secondary use of the data.

2. Through a third party

Investigators can share their data by transferring it to a data archive facility to distribute more widely to the scientific community, to maintain documentation and meet reporting requirements. Data archives are particularly attractive for investigators concerned about managing a large volume of requests for data, vetting frivolous or inappropriate requests, or providing technical assistance for users seeking to help with analyses.

3. Using a data enclave

Datasets that cannot be distributed to the general public due to confidentiality concerns, or third-party licensing or use agreements that prohibit redistribution, can be accessed through a data enclave. A data enclave provides a controlled secure environment in which eligible researchers can perform analyses using restricted data resources.

4. Through a combination of methods

Investigators may wish to share their data by a combination of the above methods or in different versions, in order to control the level of access permitted.

Timeframe for Data Sharing

As the value of data is often dependent on its timeliness Cancer Research UK expects that data sharing should occur in a timely manner. Cancer Research UK acknowledges that the investigators who generated the data have a legitimate interest in benefiting from their investment of time and effort and we therefore support the initial investigator having a reasonable period of private use of the data but not prolonged exclusive use.

Cancer Research UK expects data to be released no later than the acceptance for publication of the main findings from the final dataset (unless restrictions from third party agreements or IP protection still apply) or on a timescale in line with the procedures of the relevant research area. For example, for crystallography data there is an agreed 12-month delay between publishing the first paper on a structure and making the co-ordinates public.

With experiments carried out over an extended period of time, (e.g. population based studies), it is reasonable to expect that subsets of data analysed by the investigator(s) be made available for sharing. The investigator(s) can then continue to benefit from further reasonable periods of exclusive analysis while the dataset as a whole matures.

Research involving human participants

Investigators carrying out research involving human participants must ensure that consent is obtained to share information; furthermore the necessary legal, ethical and regulatory permissions regarding data sharing should be in place prior to disclosing any data. Every effort must be made to protect the identity of participants and, prior to sharing, data should be anonymised. In addition, any indirect identifiers that may lead to deductive disclosures should be removed to reduce the risk of identification. In most instances, sharing data should be possible without compromising the confidentiality of participants but if there are circumstances where data needs to be restricted due to the inability to protect confidentiality this should be fully addressed in the data management and sharing plan.

Data Sharing Requests

When a principal investigator is contacted with a request to share his/her data, they may ask the requestor to provide a brief research proposal on how they wish to use the data. It could include the objectives, what data are requested, timelines for use, intellectual property and publication rights etc. This may form the basis of a Data Sharing Agreement (see below). If the principal investigator has doubts over scientific validity of the proposal or the requestor's ability to analyse/interpret data correctly, this should be discussed with the requestor. A refusal to share data in such circumstances must have clear justification.

Data Sharing Agreements

To ensure that data are used appropriately investigators may consider implementing a data sharing agreement that indicates the criteria for data access and conditions for research use. This can ensure the responsibilities of both parties, along with intellectual property, citation and publication rights are agreed at the outset. It may incorporate privacy and confidentiality standards, as needed, to ensure data security at the recipient site and prohibit manipulation of data. For further guidance on managing data access and the development of data sharing agreements please refer to the 'Samples and Data for Cancer Research: Template for Access Policy Development' document available from the NCRI website: <http://www.ncri.org.uk/default.asp?s=1&p=8&ss=9>.

Data Acknowledgement

As a minimum, researchers using shared data are expected to acknowledge the investigators who generated the data upon which any published findings are based. When both parties have collaborated using a shared dataset, co-authorship on publications may be more appropriate. Researchers using shared data are also expected to acknowledge Cancer Research UK for supporting the original study.

Data Preservation

Once the funding for a project has ceased researchers should preserve all data resulting from that grant to ensure that data can be used for follow-up or new studies. Cancer Research UK expects that data be preserved and available for sharing with the science community for a minimum period of five years following the end of a research grant.