# Cancer Research UK (CRUK): CRUK Standard Template

## Data description

Outline the volume, type, content and format of the final dataset

## Data standards

State the standards that will be utilised for data collection and management

## Metadata and documentation

Outline the metadata, documentation or other supporting material that should accompany the data for it to be interpreted correctly

*Guidance*:

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).

## Data sharing

Outline the method used to share data

*Guidance*:

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:

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.

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.

Using a data enclave

Datasets that cannot be distributed to the general public due to confidentially 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.

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.

State the timescale for public release of data

*Guidance*:

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.

Explain any reasons why there may be restrictions on data sharing

*Guidance*:

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.

e.g. Development arrangements through Cancer Research Technology including intellectual property protection and commercialisation

e.g. Proprietary Data - restrictions due to collaborations with for profit organisations International policies governing the sharing of data collected outside of the UK

My research seeks supports from both the public and private sectors. How do I deal with the sharing of data? Where research is funded by a commercial sponsor, restrictions on data sharing may apply in arrangements agreed with the sponsor. Any such restriction(s) should be highlighted in the data management and sharing plan. In the event that researchers apply for or receive commercial funding for any part of their research that Cancer Research UK supports they should advise Cancer Research Technology of the situation without delay.

e.g. Confidentiality, ethical or consent issues that may arise with the use of data involving human subjects.

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.

## Preservation Plan

State the long-term preservation plan for the dataset

*Guidance*:

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.