A Guide to the Medical Research Council (MRC) Data Management Plan

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Author</th>
<th>Edits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21/03/2016</td>
<td>Gareth Cole</td>
<td>First version</td>
</tr>
<tr>
<td>2</td>
<td>17/10/2016</td>
<td>Gareth Cole</td>
<td>Edits following MRC website update</td>
</tr>
</tbody>
</table>

This document is an adaptation of guides previously created by the data.bris team at the University of Bristol.
Introduction:

In line with other research funders the Medical Research Council (MRC) expects valuable data arising from MRC-funded research to be made available to the scientific community with as few restrictions as possible so as to maximize the value of the data for research and for eventual patient and public benefit. The MRC also expects data to be properly managed throughout its life-cycle and for it to be released with appropriate high-quality metadata.

You should include a data management plan (DMP) for all research grant and fellowship proposals submitted to the MRC. This includes applications for the extension or renewal of existing funding. Additionally, for all population & patient based studies, the DMP should indicate how the study meets the requirements of the MRC's detailed guidance on data sharing for population and patient studies, particularly around access criteria and independent oversight, the means for ensuring the study and its variables are readily discoverable, and specificity about use of formal data standards.

For MRC Institutes and Units you should also develop a DMP as part of the Quinquennial Review (QQR) report.

Your data management plan (DMP) will be reviewed as an integral part of the grant or fellowship application. Reviewers will have guidance on how to evaluate a DMP.

You should keep your DMP as concise as possible and the detail should be proportionate to the complexity of the study, the types of data being managed, their anticipated long-term value and the anticipated data security requirements:

1. For population cohorts, genetic, omics and imaging data, biobanks, and other collections that are potentially a rich resource for the wider research community, a DMP may be up to 3 pages of A4.
2. Longitudinal studies, involving a series of data collections, may exceed this limit if needed.
3. Otherwise, DMPs may be as short as a quarter of a page, where the scale, complexity and costs of data management and sharing are less substantial.

What to include in your DMP?

The MRC has helpfully created a template you can use to write your DMP. If you opt not to use the template the topics listed in the template must be addressed. The template is split into seven sections:

1. Description of the data
2. Data collection / generation
3. Data management, documentation and curation
4. Data security and confidentiality of potentially disclosive personal information
5. Data sharing and access
6. Responsibilities


7. Relevant institutional, departmental or study policies on data sharing and data security

**Description of the data:**

In this section you need to outline (in no more than three lines of text) the type of study (or studies) for which the data are being collected.

You also need to outline the types (quantitative, qualitative, tissue samples, interviews, questionnaires) and format and scale of the data (e.g. file formats, software used, number of records, estimated size of the data). You should ensure that the formats and software you use enable the sharing and long-term validity of the data. Where possible, you should use open formats and technologies rather than commercial technologies. The advantage of open technologies is that they are supported by a community of users rather than a vendor or group of vendors. This means that they are less sensitive to commercial pressures and thus less likely to be withdrawn from use. The research need should always be paramount and if this requires you to use commercial technologies then you should state that here.

**Data Collection / generation:**

You don’t need to justify in your DMP why new data collection is necessary. This justification should be included in your case for support.

You do need to state how the data will be collected or generated (the methodologies for the data collection/generation) and which community data standards (if there are any) will be used.

You also need to state how the quality of the data collection will be controlled and documented. For example, outlining processes of calibration, repeat samples or measurements, data entry validation etc. In addition, if there are controlled vocabularies for your subject area then you should those where applicable.

**Data management, documentation and curation:**

You should keep this section concise and accessible to readers who are not data management experts.

*Managing, storing and curating data*

You should state how your data will be stored, backed-up, managed and curated in the short to medium term.

Loughborough University provides secure, backed up, network storage to every researcher. In addition to the Individual Workspace, Loughborough can also provide Group Workspace accessible to Loughborough University members. We would also recommend that you back up your data in at least one other place e.g. encrypted external hard drive.

---

3 [http://ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx](http://ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx)
If you are applying with researchers from other institutions then you should include in the data management plan information about the backup procedures for every Institution where data will be held.

It is recommended that you back up your data using at least two different media and, where possible, keep one copy geographically separate to the other.

If you use external hard drives or USB sticks to back up your data ensure that they are also encrypted. These devices should also be kept in a locked cupboard or office, particularly if they hold sensitive or confidential data.

If you are using personal information then you should always abide by the Data Protection Act and make this clear in your data management plan.

If you are using a laptop or mobile device then it should be encrypted so even if your laptop is stolen, other people won’t be able to access the data contained on it.

*Metadata standards and data documentation*

Sufficient metadata and data documentation must be provided to enable the dataset to be used by others. Agreed best practice standards for metadata provision should be adopted where these are in place. Metadata (“data about data”) enables other users to find and reuse your data.

The metadata will include information such as any special software you used to create the data output, what file formats is the data created in etc.

Data documentation could include the context of data collection (e.g. project history, aims, objectives and hypotheses) and data collection methods (e.g. data collection protocols, sampling design, instruments used, hardware and software used, data scale and resolution, temporal coverage and geographic coverage, and digitisation or transcription methods).

*Data Preservation strategy and standards*

If a data repository exists for your particular discipline or data type then you should offer your data to that repository. If there is not a discipline specific repository then Loughborough has created a data repository where you can preserve your data securely. This can be found at https://lboro.figshare.com. Each data record in the repository is assigned a DOI which means that it can be cited separately to your published articles, conference proceedings etc. Contact rdm@lboro.ac.uk for more information.

You should also state how long you expect your data to be preserved for. If some of your data will not be retained or preserved then you should indicate this in this section.

---

4 [http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/WTX060360.htm](http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/WTX060360.htm)
Data security and confidentiality of potentially disclosive personal information:

You only need to complete this section if your data includes personal data relating to human participants in research. The information you provide should be in line with your ethical review. You should include any formal information standards with which your study is compliant.

If your study is not using formal standards you should summarise the main risks to confidentiality and security of information relating to human participants and how these will be managed.

If you are using personal data relating to human participants then you should ensure that any device where the data is stored is encrypted. This includes laptops, tablets, external hard drives, USB sticks etc. If your data is being stored on the Workspaces (either individual or group) then you should state that this is securely (and automatically) backed up by Loughborough’s IT Services. In addition, if you are using personal data then this data should not be stored on cloud storage solutions (Dropbox, OneDrive etc.).

Data sharing and access:

You should state here whether the data you propose to collect will be suitable for sharing. If you believe it will not be then you should indicate why it is not suitable.

If your data is suitable for sharing then you need to make clear how new users will find your data. If a data repository exists for your particular discipline or data type then you should offer your data to that repository. If there is not a discipline specific repository then you may be able to use Loughborough’s data repository.

MRC’s requirement is for timely data sharing, with the understanding that a limited, defined period of exclusive use of data for primary research is reasonable according to the nature and value of the data, and that this restriction on sharing should be based on simple, clear principles. You should summarise the principles of your current (or intended) policy.

Restrictions to data sharing could include IPR, participant confidentiality or consent agreements. You should make clear in your DMP how you plan to limit such restrictions e.g. by anonymising data, gaining participant consent for sharing data etc.

If access to the research data will be restricted you should indicate who will make the decision for access to be granted. You should also highlight whether external users will be bound by data sharing agreements and set out the main responsibilities.

http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/WTX060360.htm
Responsibilities:

In this section you should specify who, alongside the PI, is responsible for data management throughout the study. This may be more than one person, particularly if the study is across a number of institutions or organisations.

Relevant institutional, departmental or study policies on data sharing and data security:


Citing Research Data in Publications:

The RCUK policy on Open Access\(^6\) requires 'all research papers, if applicable, to include a statement on how underlying research materials, such as data, samples or models, can be accessed. However, the policy does not require that the data must be made open.' The policy does not include any recommended wording for this statement. Some suggested wording could include:

- The research materials supporting this publication can be publically accessed in [insert repository/data centre/archive here] via the following persistent identifier: [insert Persistent Identifier here]. The research materials are available under a [insert licence here].
- The research materials supporting this publication have been deposited in [insert repository/data centre/archive here]. If you wish to access these research materials please contact [insert name and contact details here].
- The research materials supporting this publication can be accessed at [insert URL here].
- The research materials supporting this publication can be accessed by contacting [insert contact details here].
- Due to the confidential nature of some of the research materials supporting this publication not all of the data can be made accessible to other researchers. Please contact [insert contact details here] for more information.

Helpful Links:

MRC policy on research data sharing: http://www.mrc.ac.uk/documents/pdf/mrc-data-sharing-policy/

MRC Template for a data management plan: http://www.mrc.ac.uk/research/policies-and-resources-for-researchers/data-sharing/

MRC Guidance on data management plans: http://www.mrc.ac.uk/research/policies-and-resources-for-researchers/data-sharing/

\(^6\) http://www.rcuk.ac.uk/research/outputs/
MRC policy on sharing of research data from population and patient studies:  