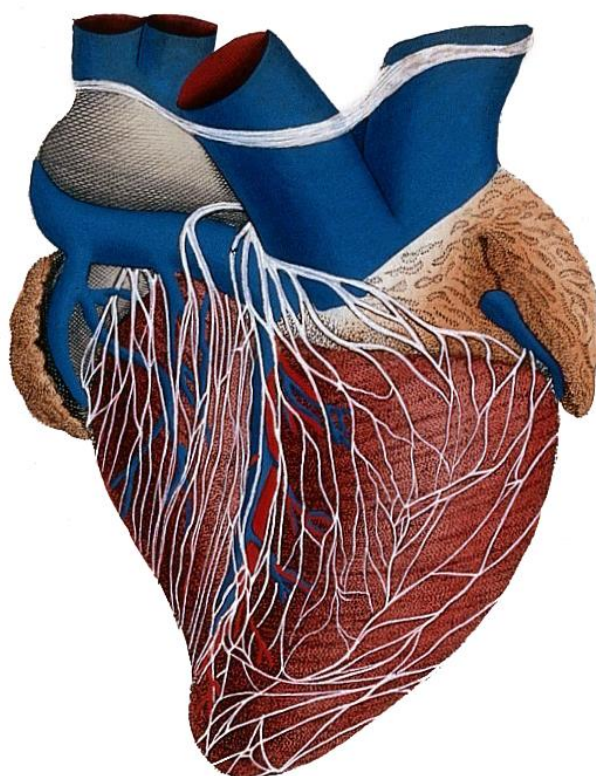Data Management Planning

Wellcome Trust funding applicants

Version 2.5 June 2021

University of Bristol

Research Data Service

Image: from Illustration of human viscera, Paulo Mascagni,

The Wellcome Library, CC-BY 4.0

SUMMARY

In common with many other major research funders, the Wellcome Trust expects researchers it funds to properly manage and, where possible, share research data. Wellcome's aim is "to maximise the availability of research data with as few restrictions as possible".[1] This applies to all researchers in receipt of or applying for funding from the Wellcome Trust's biomedical sciences and medical humanities funding streams.

Expectations of Particular Note

- All funded researchers have a responsibility to "maximise the availability of research data with as few restrictions as possible";
- Applicants are typically required to submit an outputs management plan at the application stage;
- It is the responsibility of the applicant to identify appropriate data repositories and then to deposit data into one of them to allow sharing. This should be done in a timely manner and before any research outputs which cite the data are published;
- A limited and defined period of exclusive data use is reasonable;
- Any published outputs must include a reference to underlying datasets. The default expectation is that these datasets will be available for research validation purposes. If you know that this will not

be the case, your reasons must be given within your outputs management plan.

INTRODUCTION

All those seeking Wellcome Trust funding must consider their approach for managing and sharing data at the research proposal stage. In cases where the proposed research is likely to generate data outputs that will hold significant value as a resource for the wider research community, applicants will be required to submit an 'outputs management plan' (equivalent to a data management plan) prior to an award being made.[2]

An outputs management plan should be submitted by all projects which meet any of the following criteria:

- All applications where a primary goal is to create a database resource or software tool;
- Any applications that might generate a 'community resource' as defined by the Fort Lauderdale[3] and Toronto statements[4] (i.e. large-scale, broad utility, reference dataset or dataset with community buy-in);
- Other proposals generating large-scale or other high-value data outputs with clear utility to research questions beyond those the generators of the data are seeking to address;

---

[1] Policy on data, software and materials management and sharing, https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-and-sharing
[2] How to complete an outputs management plan, https://wellcome.ac.uk/funding/guidance/how-complete-outputs-management-plan

[3] Fort Lauderdale Statement (2003), www.genome.gov/Pages/Research/WellcomeReport0303.pdf
[4] Toronto Statement (2009), www.nature.com/articles/461168a

- Any research that is expected to generate significant intellectual property (IP).

Researchers generating smaller-scale and limited data outputs or applying to the Wellcome Trust Public Engagement schemes are not required to submit an outputs management plan but are nonetheless expected to follow best practice by depositing data in an appropriate data repository in a timely manner.

In addition, the Trust requires funded publications to include a clear statement about how data and other materials underpinning the publication can be accessed in order to verify results.[5]

Your outputs management plan will be reviewed by the Trust as an integral part of the application; a poor plan can have a negative impact on an otherwise strong application. Conversely, grant reviewers will take into account any efforts made by researchers to use outputs to deliver health benefits or assist further research.

Output management plan format

The Wellcome Trust does not provide an outputs management plan template; instead they ask applicants to create a plan which is proportionate, both to the scale of the datasets generated and their likely value to the research community.

Your outputs management plan should demonstrate how you expect to fulfil your research data management responsibilities, identify any obstacles to doing so, and describe the measures you will take to meet these challenges. Your plan can also help to justify any funding required to carry out your data management activities. The Wellcome Trust requires applicants to address the following topics:

1. Data and software outputs
2. Research materials
3. Intellectual property
4. Resources required

Each topic is discussed in more detail below.

1. Data and software outputs

You should describe **what** data and software will be generated, **how**, **where** and **when** these will be shared with other researchers, and any **limits** on sharing that will be applied.

What data and software will be generated?

Consider the types of data and software that will be created, and which of these will have re-use value to other researchers. You should also describe any formats and standards that you will use to maximise effective sharing.

In attempting to organise and document your data, it may help to imagine a secondary data user trying to make sense of your data in your absence, after your project has concluded. If no metadata or other documentation were provided, this secondary user would be faced with the difficult task of 'unpicking' your data. How, for instance, would they make sense of your file and folder names? Or your methodology or approach to data processing? What extra information

---

[5]Requirements for publishing Wellcome Trust-funded research papers:

wellcome.ac.uk/sites/default/files/requirements-publishing-wellcome-funded-research-papers.pdf

would they need to make the most of your data? Consider whether it is appropriate to use an existing disciplinary metadata standard[6] to help you organise your data.

The Trust recognises that in some cases it may not be appropriate for researchers to share their data. However, if your research meets the criteria for requiring an outputs management plan outlined above, but you plan not to share your data, your reasons must be clearly explained within your plan.

How will data and software be preserved?

Your outputs management plan should describe how datasets that have long-term value will be preserved and curated beyond the lifetime of the grant.

You should explain where your data will be stored, how it will be organised in the short term and who will back it up. If you are not part of a study with existing data storage arrangements, it is recommended that you store any data you create in the University's Research Data Storage Facility (RDSF) managed by the Advanced Computing Research Centre (ACRC).[7]

The back-up procedures, policies and controlled access arrangements used by the RDSF are of a very high standard. Each research staff member is entitled to 5TB of secure data storage without charge, and any data stored will be preserved for a minimum of 20 years. If your storage quota is already used up, or if your project will exceed this storage limit, there will be a cost and ACRC should be contacted for guidance before your budget is finalised.

If you do not intend to make use of the RDSF, your storage provider's back-up and security procedures should be described within your Plan instead. If you will be working collaboratively with other institutions, make sure that the security and back-up procedures of each data-holding partner are described.

However, if you intend to keep data on a local (e.g. departmental or research-group based) server, you should ensure that you have the resources and systems in place so that the data will not only be safely stored, but will continue to be curated, secured and shared in a way that maximises its value and addresses any associated risks. This includes consideration of how data held in this way can be effectively linked and integrated with other, external datasets to enhance its value, and what will happen to the data in the long term.

Where will shared data and software be made available?

Unless there is a compelling reason not to do so, researchers are expected to deposit data in recognised data repositories and the Wellcome Trust maintains a list of major data repositories including biomedical repositories that preserve and provide access to research data.[8] Additionally, the University of Bristol has its own research data repository providing several different levels of access to data which researchers from any discipline may wish to use. Access options

---

[6] DCC list of disciplinary metadata standards: http://www.dcc.ac.uk/resources/metadata-standards
[7] Advanced Computing Research Centre, University of Bristol, www.acrc.bris.ac.uk

[8] Wellcome Open Research approved data repositories https://wellcomeopenresearch.org/for-authors/data-guidelines#hosting

range from entirely open to rigorously controlled, which is suited to 'sensitive' data.[9]

When will data and software be shared?

Timescales for sharing data should take account of any standards of good practice in the applicant's research field (for example Fort Lauderdale Principles[10] and subsequent Toronto statement[11] on pre-publication data sharing). The Trust recognises that data producers have the right to a reasonable, but not unlimited, period of exclusive use of the research data which they have produced.

Where appropriate, researchers may use publication moratoria to facilitate pre-publication sharing of data with other researchers, while protecting their right to first publication. Any such restrictions on data use should be reasonable, transparent and in line with established best practice. The RED contracts team can advise on any formal moratoria you wish to put in place.[12]

As an absolute minimum, data underpinning research papers must be made available to other researchers on publication, providing this is consistent with any ethics approvals and consents which cover the data.

Limits to data sharing

Where a managed access process is required, for example, where a research activity involves potentially identifiable data about research participants, researchers must use appropriate anonymization and

the access mechanisms should be proportionate to the risks associated with the data and must not unduly restrict or delay access. Access procedures must be consistent and transparent and must be described as part of your outputs management plan.

2. Research materials

This section should address the management of materials such as antibodies, cell lines and reagents that are developed during the course of your research. You should describe how these materials will be made available to other users, for example by submitting to a recognised biobank such as the European Collection of Authenticated Cell Cultures[13] or by licensing them to a suitable business partner who can handle storage and distribution.

3. Intellectual property

You should describe what IP your research will generate, how this IP will be protected, and how IP will be used to maximise health benefits. The Wellcome Trust recognises that open sharing is not always appropriate, and that commercialising your outputs may allow for further development and distribution to maximise adoption and use by the wider research community. This may be particularly appropriate for software and research materials such as cell lines.

In this section, you should set out how you intend to identify and capture predicted IP as well as unanticipated discoveries, and briefly describe how

---

[9] data.bris repository, https://data.bris.ac.uk/data/
[10] Fort Lauderdale Statement (2003), www.genome.gov/Pages/Research/WellcomeReport0303.pdf

[11] Toronto Statement (2009), https://www.nature.com/articles/461168a
[12] https://www.bristol.ac.uk/red/contracts/
[13] ECACC: https://www.phe-culturecollections.org.uk/collections/ecacc.aspx

you intend to protect this IP (for example, by applying for patents). The RED commercialisation team[14] can advise on methods and timescales for protecting IP.

The Wellcome Trust regards health benefit as the primary driver for any commercialisation, so you should explain how your chosen IP strategy will bring benefits to the wider research community.

4. Resources required

You should consider carefully any resources you may need to undertake data management and sharing. Where dedicated resources are required, these should be outlined and justified as part of your plan. The Trust suggests applicants consider:

- People and skills: is there sufficient expertise and resource within your research team to manage and share your data effectively? Is additional specialist expertise (or training for existing staff) required?
- Storage and computation: do you have access to computational facilities required to manage, store and analyse data generated by your research? Will additional computational resources be needed in order to process or analyse your data, and what will be the costs associated with these?
- Access: do you need to operate a data access committee to manage requests for data over the lifetime of the grant? If your research is using

third party data, are there costs associated with accessing it?

- Preservation: does your chosen repository charge for depositing data? Note that costs should be incurred within the lifetime of the grant.

Support at Bristol

Several support services are in place within the University to help you manage and share your research data. These services include: ACRC (Research Data Storage Facility),[15] your Faculty IT team (general IT support),[16] the Research Data Service (research data management training and guidance),[17] RED (advice on collaborative projects, contracts and IPR)[18] and the Office of the Secretary (for Data Protection and FOI).[19]

---

[14] http://www.bristol.ac.uk/business/innovate-and-grow/research-commercialisation/contact-us/
[15] https://www.acrc.bris.ac.uk/storage.htm
[16] https://uob.sharepoint.com/sites/itservices/SitePages/fits.aspx

[17] https://www.bristol.ac.uk/staff/researchers/data/contacts
[18] https://www.bristol.ac.uk/red
[19] https://www.bristol.ac.uk/secretary