Guidance from the Early Career Researcher Advisory Board

Wellcome Open Research

Things I wish I had considered at the outset of sharing my imaging data

Ben Steventon

"In hindsight, there are two major stumbling blocks I have come across when it comes to the open sharing of my imaging data. The first relates to data annotation and is a major issue when it comes to ensuring image data sharing complies with FAIR principles. Each imaging experiment has many variables, from the biological specimen through to method of sample labelling, choice of microscope, imaging parameters and subsequent image processing and analysis.

The problem is further compounded by the fact that each imaging experiment may not involve the same set of variables, meaning the image annotation itself varies from one set of experiments to another. Therefore, I wish I had considered the use of a flexible database through which to store all the labs images, that can be annotated or tagged in different ways as the research progressed. We are now using Omero for this, but are currently presented with a back-log of work in terms of adding our images to the database. We will get there though!

The second issue comes to the choice of image data repository to use. Again, there is a lot of variation in the best choice for this from experiment to experiment. The main issue here is data-set size, with our images ranging from a few hundred megabytes up to several terabytes. As such, there is no 'one place fits all' solution in terms of image data repository, and each has their own restrictions and ways to manage the uploading of images, annotations and metadata. Unfortunately, we have not found a solution for this as yet, and are currently developing ways to host and share our data openly with the help of a Wellcome Trust Open Research Enrichment grant. So, more on this soon."

Wellcome Open Research

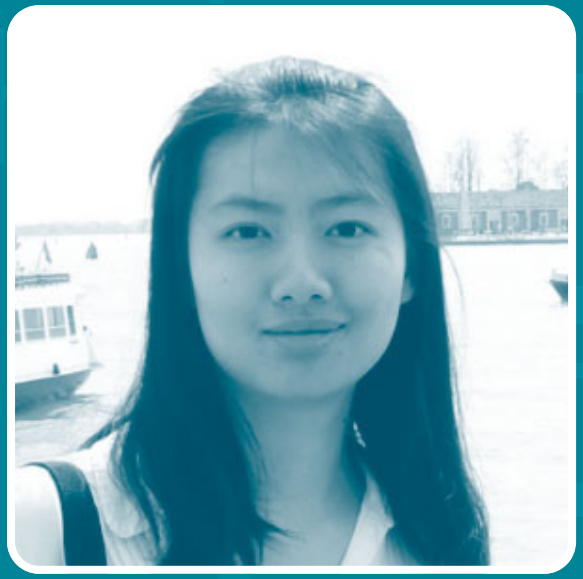
Think about data sharing when you write your informed consent forms

Fiona Cresswell

"As a clinical researcher, I recently spent a lot of time designing the participant information leaflet and informed consent forms for the clinical trial that was part of my PhD research. I was keen to make the research easy to understand, yet sufficiently detailed to accurately represent all the procedures, risks and benefits involved. Now that the study has been completed, I realise that I did not include enough information about data sharing on the informed consent forms.

At the outset it would have been ideal to seek consent to share their anonymised data on an open-source repository. I will now need to revisit the Research Ethics Committee and possibly even the participants themselves to seek this approval retrospectively. I will not make this mistake again as going back to the regulatory bodies at the end of a study is not my idea of fun! I hope this tale will help you to avoid the same pothole in the data sharing journey."

Wellcome Open Research



Things I would do to prepare for data sharing in a new study

Cherry Lim

"As an epidemiologist, I often work with clinical data, which can be sensitive to share. There are three major things I would do when planning for a future study in preparing to share the collected data. Firstly, I would expose myself to guidelines on data sharing. These guidelines include local regulations, funder guidelines/requirements, and tutorials on data sharing. This is important especially with clinical data, which often involves study participants' confidential information. I would list the can and cannot share information to guide my data management plan.

Secondly, I would discuss and communicate with the study site on my plan for data sharing as early as possible. Sharing the findings and data with the local study site on a regular basis has helped me in the past in gaining trust from the local site. I had meetings to report the findings every three months in the first year of my study, and I prepared cleaned data that can be readily shared with my study site. Through this process, I had not only gained trust from the local study site, but also generated opportunities for further collaborations with the local site. We are now planning a new study of an idea that came from the hospital staff based on the data I have collected and shared with them.

Finally, I would communicate with the data sharing committee in my host research institution as early as the stage of drafting protocol. Seeking expert guidance and support on how to share, what to prepare for, and which resources to utilise as early as possible can save a lot of time and effort."

Wellcome Open Research

Considerations I would prepare at the beginning, or actually even before, any new study

Jana Hutter

"As a researcher working at the interface of novel data acquisition strategies and clinical questions, ethical considerations and the preparation of the corresponding ethics application are key and time-consuming requirements for all of my research.

Despite the time invested, I have in the past been in the situation that I wanted to share my data, but had not included the right wording in the participant information sheet and informed consent form, a source of much frustration!

If I had carefully considered all the details and consulted with all partners at the very beginning on both the requirements and the opportunities for open data sharing, it would have saved me a number of amendments and regrets later on! I would thus, for any new study, take the time required at the very beginning and write down in detail which type of data exactly (e.g. imaging data, clinical data, cohort data, supporting data such as technical scanner information) could be useful for stakeholders. Furthermore, I would then discuss these thoughts and considerations during the corresponding meetings with PPI groups (Participant Patient Involvement groups) to ensure that the wording and information given addresses the needs and requirements, but also most importantly, any possible questions and worries the participants would have."

Wellcome Open Research

The benefit of hindsight

Rebecca Payne

“I work with clinical research data sets as well as lab-based data and clearer wording around how anonymised data can be shared during informed consent would have saved a lot of time and anxiety! I would also have engaged with local facilities managers, data stewards and data managers at my Institute much earlier to seek advice about data management. I work with large datasets and wasn’t quite prepared for quite how large these would become.

Good luck.”

Wellcome Open Research

Details I would consider in a Data Management Plan for a grant application

Mohlopheni Jackson Marakalala

“I work on projects that tend to generate clinical/proteomics datasets. However, I used to struggle on the exact details to include in a data management plan, especially when applying for grants. With consultation and the little experience I’ve gained over time, I would recommend the following points:

- What is the nature of the data outputs that will be generated by the research and what gaps it will address in the field?
- Availability of the data: When will the data be made widely available to user communities?
- Accessibility: where will the data be deposited and whether the public data repository site is easily accessible to target audience. I find it exciting that Wellcome Open Research has created a platform to publish articles in ‘data note’ format, providing an opportunity for researchers to share their resources on an open access platform. ”

Wellcome Open Research