Summary of the NCRI Data Sharing Meeting

Cancer Research UK, Holborn, London

Tuesday 6th July 2010

Background

Since its formation the NCRI Informatics Initiative has endeavoured to address the issues surrounding data sharing and integration in cancer research. In 2003, an NCRI Data Sharing Group was formed which subsequently developed a data sharing policy that was formally adopted by the NCRI Board in July 2005.

The NCRI Data Sharing Policy[1] served as the first step in encouraging wider access to data generated in cancer research by encouraging the NCRI organisations to introduce a requirement for scientists to include data sharing strategies in funding applications and for provision to be made by funders to support the cost of data sharing where necessary.

Since the adoption of the data sharing policy the NCRI Informatics Initiative has been working with the NCRI Partner organisations to implement data sharing through developing individual data sharing policies and guidelines that set a clear direction to allow scientists to carefully consider data sharing when applying for funding

Currently the NCRI member organisations are at different stages of the implementation process and the NCRI Data Sharing Meeting was organised to bring together as many of the NCRI Partner organisations as possible to provide a mechanism to discuss the challenges that the NCRI Partners are experiencing in implementing their individual data sharing policies and how these challenges can be addressed.

Meeting Summary

Prior to the meeting, specific areas were identified as being particularly challenging and these became the focal points for discussion. These areas included:

- Compliance and Review of Data Sharing Policies;
- Guidance and Training;
- Ethics, Consent, Confidentiality and IP;
- Financial Issues; and
- Data Storage and Access.

Compliance and Review of Data Sharing Policies

Many funders now require scientists applying for funding to submit a data sharing plan that details how they intend to make their data accessible to the wider research community. In most cases the data sharing plans are reviewed as part of the funding decision. The NCRI Data Sharing Meeting identified a number of issues associated with the introduction of data sharing to the grant application process. These issues include the variability in the quality of data sharing plans and the lack of expertise of reviewers whose responsibility it is to evaluate the plans. In addition, there are questions about how to monitor

whether data sharing plans are being complied with. Organisations like the BBSRC and Wellcome Trust have started to address some of these issues. For example the BBSRC have introduced training for their reviewers to ensure that they have the skills needed to properly evaluate data sharing plans and they have a Data Sharing Monitoring Group to assess compliance. The Wellcome Trust are in the process of developing clear guidelines to assist scientists in the preparation of data sharing plans and are developing a Code of Conduct for data sharing in the field of Public Health.
Additionally, it was felt that underpinning the issue of compliance with funders' data sharing policies is the lack of incentives to share data. There was a feeling that significant incentives are needed to encourage data sharing. It was suggested that these incentives be centred not only on funding but demonstrating very clearly the benefits of data sharing. Furthermore linking data sharing to career progression was felt to be an effective mechanism to encourage data sharing and the BBSRC are exploring this through determining the possibility of making a scientist's shared dataset part of their career record.

Guidance and Training

It was determined that most organisations provide some form of guidance to their researchers but that in some instances this guidance was broad and needed to be clearer to ensure that there is a good understanding of what is expected of researchers when they are developing a data sharing strategy. It was also identified that it would be beneficial to introduce training for reviewers as the BBSRC has done as well as short training courses for researchers in data sharing, management and curation. It was felt that some form of training in this area could be introduced during the induction days that a number of funders hold for their grantees.

Ethics, Consent, Confidentiality and IP

It was felt that this area was very straightforward in that there is clear guidance on the sharing of personal, sensitive or confidential data. This type of data can be shared provided informed consent is given and if it is not then researchers either cannot share this data or can apply to share data under the power of Section 251 of the NHS Act 2006. Here, scientists have to demonstrate that the benefit to the public outweighs the rights of participants' to privacy and they also have to show that obtaining consent was not practicable.

In terms of IP issues there was concern over how data generated from industry/academia partnerships could be shared without breaching IP but it was highlighted that this can be addressed early on in the planning of a project with permission granted sought from all data owners on what data can be shared and when.

Financial Issues

All funders in attendance stated that the cost of data sharing was an integral part of the grant application process and researchers could factor in any data sharing costs. Funders had observed that researchers were not requesting funding for data sharing and felt this could be because they were unsure of the resources required for data sharing both in terms of time and cost. Training and guidance could provide researchers with the information on which to conduct a proper cost assessment of their data sharing requirements.

Data Storage and Access

It was felt that there are a several options in terms of technical infrastructure to support data sharing in the UK. It was mentioned that a review of the UK's e-infrastructure was recently conducted outlining the requirements for the UK and the results of this would be published in September of this year. It was felt this would provide an idea of any gaps that might exist.

It was stated that the BBSRC have a funding stream for the development of resources to support data sharing and that CRUK are working with SAGE, a project that aims to develop standards-based infrastructure to support data sharing. A number of CRUK scientists have made use of resources from SAGE and where there are any gaps in infrastructure CRUK themselves provide the means to fill these gaps. An example given was in the development of infrastructure to connect the Experimental Cancer Medicine Centres. The technical infrastructure to support data sharing is very important to encourage data sharing and funders are encouraged, that where relevant and possible, they provide the resources to initiate the development of such infrastructures

Recommendations

Based on this meeting the NCRI Informatics Initiative recommends the following:

- Clear guidance should be given to assist applicants in developing their data sharing plans
- There should be consistency across the different funders with respect to the requirements for data sharing plans
- Some form of training for reviewers should be considered
- Examples of good data sharing plans should be compiled and shared to help guide applicants
- Specific research examples of the benefits of data sharing should be collated to assist in encouraging data sharing