Data management and sharing: a funder perspective

LSHTM meeting on developing data management expertise in research
6 November 2012

David Carr
Wellcome Trust
d.carr@wellcome.ac.uk
Sharing research outputs: our position

• committed as a funder to ensuring research outputs (inc papers & data) are made available in a way that maximises benefits to health and society

• we are a leading advocate of open access & data sharing:
  – long-standing policies on open access publishing & data management and sharing
  – funding to implement policies and support for key enabling infrastructures
  – work in partnership at UK and international level to provide enabling environment
Data sharing – a growing consensus

• major challenges associated with increasingly vast & complex datasets, but also tremendous opportunities

• increasing policy convergence between major funders in promoting sharing of research data
  – expectation that data outputs be preserved and shared in a way that maximises value
  – requirement for data management plans as integral part of the application process

• but implementation and enforcement of these policies is still rather variable
The Wellcome Trust’s policy

• Data management and sharing policy originally published in 2007, updated in 2010

• Key elements of our policy:
  – all funded researchers should maximise availability of data with as few restrictions as possible
  – data management and sharing plans to be provided as part of applications, where research likely to generate data of value as a resource for the wider community
  – data management and sharing plans and associated costs considered as integral part of funding decision.
Policy implementation – our experience

• we reviewed implementation of our policy in 2010, and introduced new guidance – setting out key expectations for data management and sharing plans

• quality of plans has been variable, but good evidence that they are improving

• Challenges remain:
  – post-award tracking of plans
  – awareness of staff and reviewers
  – are resource implications of plans assessed appropriately?
Major challenges

• there are significant barriers & constraints to overcome:
  – Infrastructural issues
  – Cultural issues
  – Technical issues
  – Professional issues
  – Ethical issues

• different disciplines at very different stages; different types of data raise distinct issues

• challenges will require funders to work in partnership, with each other & other key communities
Working in partnership: three current initiatives

- **ELIXIR** – sustainable infrastructure for European life science data

- **Expert Advisory Group on Data Access** – advice on emerging issues relating to data access across genetics, epidemiology and social sciences

- **Public Heath Research Data Forum** – cross funder initiative to increase access to research data generated by public health and epidemiology research
Increasing access to public health and epidemiology research data

• current situation is often that:
  – lack of access to data limits comparative analysis or even checking for accuracy
  – only immediate research interests are addressed; broader questions are neglected
  – lots of duplication in data collection
  – data not archived effectively for long-term

• recognise there are legitimate concerns, and that limits and controls are essential…

• … but growing recognition among funders that we need to find ways of widening access to this data
Sharing research data to improve public health

• Joint statement of purpose was published in January 2011, with 17 initial funder signatories

• the statement set out:

  – **one vision** – to increase availability of research data generated by our funding to accelerate improvements in public health

  – **three principles** – share data in ways that are **equitable**, **ethical** and **efficient**

  – **seven goals** – immediate goals and longer term aspirations
Three Principles

**Equitable:**
Any approach to the sharing of data should recognize and balance the needs of researchers who generate and use data, other analysts who may want to reuse those data, and communities and funders who expect health benefits to arise from research.

**Ethical:**
All data sharing should protect the privacy of individuals and the dignity of communities, while simultaneously respecting the imperative to improve public health through the most productive use of data.

**Efficient:**
Any approach to data sharing should improve the quality and value of research and increase its contribution to improving public health. Approaches should be proportionate and build on existing practice and reduce unnecessary duplication and competition.
Seven goals

Immediate goals
1. Data management standards support data sharing
2. Data sharing is recognized as a professional achievement
3. Secondary data users respect the rights of producers and add value to the data they use

Longer-term goals
4. Well documented data sets are available for secondary analysis
5. Capacity to manage and analyse data is strengthened
6. Published work and data are linked and archived
7. Data sharing is sustainably resourced for the long term
A broad range of funder partners

Signatories...

Supporting organisations...
Implementing the joint statement

• partner funders have formed a **Public Health Research Data Forum**, with a joint workplan

• activities in three main areas:
  – capacity and skills
  – culture and incentives
  – infrastructure and tools

• challenges will require key inputs from researchers, institutions, funders, repositories and publishers
Forum activities

• creating a resource of funder policies
• developing options for capacity building and training in LMICs
• commissioning work on data citation mechanisms
• building the evidence base on the perspectives of research participants and stakeholders in LMICs
• developing accessible guidance on metadata and standards
• scoping a potential workshop with African academies on data sharing
Some closing thoughts

• funders and institutions each have critical roles to play in building the culture and resources needed to maximise value of research data

• importance of institutional planning and coordination:
  – assess resource needs for current and future data assets
  – articulate clear policies and processes to support researchers
  – robust evaluation and monitoring

• particular challenges of working with LMICs - equity and sustainability
Further information

http://www.wellcome.ac.uk/publichealthdata

d.carr@wellcome.ac.uk