

**MINUTES OF THE SECOND MEETING OF THE EXPERT ADVISORY GROUP ON DATA
ACCESS (EAGDA)**

10.45-16:30, FRIDAY 19 OCTOBER 2012, WELLCOME TRUST

Present:

James Banks
Martin Bobrow (Chair)
George Davey Smith
Rosalind Eeles
Mark Guyer
Tim Hubbard (via telephone)
Mark McCarthy

Andrew Morris
Bartha Knoppers
Onora O'Neill
Chris Skinner
Melanie Wright

Geraldine Clement-Stoneham- MRC

Katherine Littler – Wellcome Trust
Priya Umachandran Wellcome Trust

Apologies for absence:

Paul Burton
Peter Dukes – MRC
John Hobcraft – ESRC
Fiona Reddington - CRUK

Peter Elias- ESRC
Suzanne Mills ESRC
Nigel Shadbolt

1. Membership

- The European Genome Archive has been contacted about co-opting an EAGDA member.

2. Workplan

- The Group discussed the papers prepared by the Secretariat on the themes identified at EAGDA's first meeting. They were:
 - Ownership and control of data
 - Establishing incentives and changing cultures to support data sharing
 - Governance and mechanisms for access
 - Quantifying Risk
 - Data Access Committees
- The Chair outlined proposals for establishing smaller working groups, to tackle the themes the Group will take forward.
- The Group was reminded that resources were available for commissioning research, if deemed necessary.
- The Group discussed possible outcomes for EAGDA work including toolkits, guidelines, recommendations, advice but concluded that the outcome would be relevant to the specific topic so EAGDA should keep its options open.
- It was agreed that those who attended, or who were invited to attend, the stakeholder workshop should be notified of the workplan. This would provide an additional opportunity to identify any gaps in the workplan.
- The need to make EAGDA's work as transparent as possible was reiterated, with outputs available not only to the research community but also to the wider public.

3. Ownership and control of data

- The paper, as mapped out by the Secretariat, included an analysis of access policies, consent forms and data access and material transfer agreements from several major studies and biobanks. It summarised the use of language, the effects of data linkage on control and issues of participant consent for use of the data they provide.
- The Group agreed that this would be better placed as a context paper, but that this did not preclude recommendations being made in relation to standardising language around data access.
- In terms of language, the Group noted the following:
 - The term 'ownership' was problematic and should be replaced with the term 'custodian'.
 - The paper must also make reference to the term 'stewardship'¹
- In terms of the section on control over data, the paper needs to also cover:
 - The complex role of the data controller
 - Legal liability for breach of data confidentiality – including where this stops?
 - Is there a meaningful definition of the differences between a data set and a study?
- It was suggested that the revised version of this paper should contain case studies featuring multiple data controllers, or linkages of data sets, and should examine how these systems work especially when a data breach occurs.

SUMMARY: This paper will be turned into a context paper, to include: an overview of what is meant by data and other terminology; an analysis of differences in the treatment of large and smaller data sets; a summary on legal responsibility for datasets; discussion on proportionality and an examination of the role of the data controller. Case studies will be used to help illustrate the issues.

4. Establishing Incentives and Changing Cultures to Support Data Access

- The paper produced by the Secretariat explored the role of funder policies in increasing data sharing while recognising the rights of data generators, emerging systems for measuring and recognising the use of shared data and the need to develop the skills and careers of data managers.
- A number of amendments were suggested to this paper:
 - 'Data sharing' should be dropped from the title as it covers a multiplicity of things and suggests mere disclosure, rather than purposeful and controlled contribution to a research project. This will be amended to 'data access.'
 - The paper should also cover the mechanisms and infrastructures that act as a disincentive to providing accessible data.
 - The costs to researchers of releasing data, both in terms of reputational and financial costs and the costs of obtaining the necessary ethics approvals, need to be identified. Cost recovery models e.g. Biobank were discussed, but what is unknown is how many potential users are put off by this cost.

¹ Stewardship was used in the Nuffield report from 2002 'Learning from experience: privacy and the secondary use of data in health research' by William Lowrance. [Privacy, Confidentiality, and Health Research (Cambridge Bioethics and Law)

- The Group discussed researcher recognition. In Scottish law there is protection from releasing data for a limited period. The House of Commons Justice Select Committee, chaired by Alan Beith, produced a report on EU data protection proposals in November 2012. The report recommended the same protections should be added in English law for research data.
- Any analysis of incentives needs to consider the costs/benefits to all involved in the process. What are the benefits for funders, data creators, recipients? What are the (potential) costs for data creator i.e. reputational and financial? What can be done to mitigate these costs?
- The Group also raised the issue of non-use of data, and the fact that there was no penalty involved. People can request data and then not use it, and it is a condition of funding that the request is met. Should there be penalties for non-use of data? How do you distinguish between someone who does not use data because it is not useful (in good faith) over someone who is simply 'fishing' for data?

SUMMARY: This paper will build on the discussion paper and in addition will discuss cost recovery models; non-use of data; and clarify the incentives for the different parties involved in communicating data. The Group agreed this was an area where EAGDA could usefully add value.

5. Governance and Mechanisms for Access (formerly known as Gateways of Access)

- The paper that informed this discussion examined issues around data discoverability via gateways, mechanisms of data access and the usability of accessed data.
- The Group agreed to amend the title of this topic, as it was deemed too narrow given the more general access issues this paper should cover.
- Several additions to this paper were deemed necessary:
 - The role of ethics and privacy reviews was missing in the discussion paper.
 - The national requirements for independent ethics reviews can vary from country to country and some sort of analysis should be included.
 - The issue of whether it is either possible or advisable to harmonise mechanisms of data access should be addressed
 - Is it possible to decide upon principles or tools that work from convergences between disciplines? It is important to acknowledge that not only are there differences between disciplines and the way they distribute data but there are also differences within studies in the same discipline.
- The Group note that the Cabinet Office Open Data White Paper seemed to conclude that the availability of biomedical and social data could be narrowed with certain conditions and the use of safe havens. The white paper also deals with enabling the corporate sector to use public data.
- The Group discussed a number of issues regarding Freedom of Information (FOI) Act and the need to include under this heading.
 - The legal obligation contained in the Protection of Freedom Bill, to make data re-usable in digital form, is a very serious demand.

- The Information Commissioner's Office does not always recognise the costs of making data re-usable, especially with large datasets requiring specialist software.
- FOI implies the unconditional availability of the data. Those who acquire data with no conditions have no obligations.
- It should not be assumed that FOI is not relevant for the scientific community, as it is a potentially cheaper way to access data and even access pre-publication data and could prove a disincentive for researchers to have data held in a way that can be discovered.
- Further to this, the Group discussed the formats in which data could be provided. Given that some datasets are very complex, whose responsibility is it to make them manageable and transferable? How far should any duty extend?
- Generic Access Agreements must also be considered in this context
- Following the discussion, the Group thought this might be an appropriate area in which to consider the possibility of formulating guidance or overarching principles.

SUMMARY: It was agreed that the working group looking at this topic will need to expand this paper to cover the role of ethics and privacy reviews; implications of FOI; and the imperative to provide 'useable' or 'easy to use' data. After deliberations, the Group will consider whether developing guidance of the formulation of principles is an appropriate outcome.

6. QUANTIFYING RISKS

- The paper explored the importance of quantifying the actual risks associated with data sharing, as opposed to the perceived risks. It also explored the associated ideas of risk stratification, and public perception and risk. The paper mentioned, but did not discuss, the issue of quantifying the risk of not sharing data which was felt by EAGDA to be neglected but important.
- There should also be a subsection on risk in each of the other papers, as well as in this overarching paper.
- The Group felt the paper would benefit from the following additions:
 - Important to identify cases of risk
 - Proportionality is crucial in discussions of risk and data access
 - There is a need to include examples of the processes which minimise risk
 - It will also be important to assess when and why breaches have occurred, and address the points at which the breach has occurred
 - The paper could give some sort of baseline of risk by comparing risks involved with research data to other scenarios e.g. corporate data collection and associated risks
- It was noted that this could be one of the areas where additional research may need to be commissioned

SUMMARY: The working group on this topic should put together a paper on risk discussing: who/what is at risk; risk benefits analysis, risk assessment, risk stratification and risk minimisation; dispelling myths around risk and contextualising risks; and exploring how to quantify the risks of not allowing data to be accessible. EAGDA considered this would be a challenging area to work on.

7. DATA ACCESS COMMITTEES (DACs)

- The paper on DACs highlighted the increase in the number of datasets and corresponding DACs. These DACs have emerged with little guidance on how

they should operate. Consequently, there is a varied landscape which could diminish the ease of future research, especially involving the linking of multiple datasets or international use.

- The Group felt that EAGDA could make a particular impact in this area, especially given there was less literature around DACS than other data access topics.
- The Group noted the lack of independent advice or guidelines for how DACS should operate and that this is leading to a piecemeal approach. This is something EAGDA is well placed to address.
- Issues that should be included in the paper included:
 - Membership of DACS, especially whether Principal Investigators should be DAC members
 - Should there be harmonisation of policies?
 - Is there a role for better co-ordination of DACs?
 - Categorisation of data – open vs. controlled – are these categories appropriate?
 - The growing numbers of DAC requests and whether this is problematic
 - Is there a need for training for DAC members?
 - Are there benefits in new approaches to DACS, e.g. having fewer dedicated committees to manage data access?
 - Differences between data access and sample access committees need to be outlined.
- It would be useful to consult a sample of DACS and identify what difficulties they have, building on the background research already contained in this paper and the work commissioned by the Wellcome Trust.

SUMMARY: EAGDA felt this was an area in which it could make an impact in a relatively short time scale. The follow-up paper will include an analysis of current principles and practises, highlighting key similarities and differences. The working group could then look to see where policies could be changed, or where they are working but the absence of guidance means others are not well placed to learn from these experiences. Case studies would be useful, as could the development of a toolkit.

8. Cross cutting topics

- The group identified a number of cross-cutting areas:
 - **International Dimensions**
 - There is a patchwork of existing legislation and it is difficult for researchers to see a way through this
 - Legal restrictions and the potential impact on data transfer
 - Lack of legal sanctions for breach of data sharing conditions in some jurisdictions. There is a need to explore other types of sanctions, including those set by funders
 - Important across all topics is an understanding of how data is or is not transferred, what is working and what is not.
 - **Public engagement** – the Group needs to consider how it makes its deliberations public in a useful way. This, and other aspects of public engagement, should be explored further in a future meeting.
 - **Research subject Consent** – particularly in relation to linking datasets.
 - **Liability** for data breaches
 - Who should be liable? Should it be shared between data collectors, data holding institutions and data users?
 - **Reputational risks** for data collectors and their institutions when making data accessible to external users
 - **Responsibilities** of data users and how these are enforced

- **Data Regulation**
 - The complexity of data regulation
 - Problems arising from rules widely framed, so do not apply specifically to research data
- **Data Production** – the Group reminded itself that it in considering the issues and any potential advice, it is important to try to consider the issues at the time of data creation.

9. Next steps

Working Group allocation (Martin Bobrow to participate in all groups)

Incentives

Mark Guyer
Tim Hubbard
Bartha Knoppers
(sanctions)
Melanie Wright

Risks

James Banks
(consult)
Mark McCarthy
Chris Skinner
Melanie Wright

Mechanisms

George Davey-
Smith (access
mechanisms)
Bartha Knoppers
Melanie Wright
James Banks

DAC

Paul Burton
George Davey
Smith
Mark Guyer
Chris Skinner
Ros Eeles

ACTION: The Secretariat will contact working group members in due course.

Date of next meeting:

The 7th December meeting is cancelled.

The next meeting of the whole Group will be 4th March 2013.

10. Data Protection Act Regulation

- The Group also discussed the implications of FOI Act and the EU Data Protection Regulation.
- The DPA Regulations have a number of associated issues:
 - Lack of clarity of role of data controllers
 - Demand for explicit consent, which is worrying as it stands as it could undermine longitudinal studies. Exemptions to this demand should be clarified
 - Genetic exceptionalism is reinforced in the Regulation.
 - The rights described seemed unrealistic in practical terms e.g. right to be forgotten
 - Some of the content of the Regulation is broad in its scope. Provisions designed to target social media will therefore apply to all sectors with a potential deleterious effect.
 - An additional issue is that the DPA is a Regulation not a Directive, so will not need transposition into national law. Consequently, this is a crucial time in the process. There may not be another opportunity to influence the Regulation.
- Ministry of Justice is setting up a task force to discuss this, which the Wellcome Trust will be involved in. Some of the Group questioned the role of the Ministry of Justice in leading the UK's discussion but it is too late for this to be changed.

- ACTION: The first meeting of the task force is in November, WT will feedback to EAGDA
 - The Regulations could have a huge impact on international data transfer and endanger the work of international consortia.
 - ACTION: Draft a short concise letter to funders to outline EAGDA's concerns and to highlight the need to engage government CSAs in the discussions.
- 11. FOI**
- The impact of FOI legislation and the changes proposed in the Protection of Freedoms Bill was discussed, for universities and funders. The private sector is exempt from FOI.
 - In the US, Shelby amendments require the release of all data that have been used to make government decisions.
- 12.** The Group discussed the recently launched E-health initiative, which seeks to link social and economic data with health records in the UK.