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

10:30 – 16:00 MONDAY 20 APRIL 2015, WELLCOME TRUST

Present:
Martin Bobrow (chair)
Paul Burton (until 13:00)
Mark Elliot
Paul Flicek
Mark Guyer
Tim Hubbard
John Hobcraft
Andrew Morris
Onora O'Neill
Melanie Wright

Funders:
Fiona Reddington (CRUK)
Vanessa Cuthill (ESRC)
Clare Feary (ESRC)
Jon Fistein (MRC)
Katherine Littler (Wellcome Trust)
Natalie Banner (Wellcome Trust)

Apologies for absence:
James Banks
Ros Eeles
Bartha Maria Knoppers
Mark McCarthy
Geraldine Clement-Stoneham (MRC)

1. Welcome

The minutes of the sixth meeting of EAGDA, dated 16 October 2014, were ratified.

Mark Elliot (Manchester University) was welcomed and provided a brief overview of his research interests, which include statistical disclosure control, confidentiality and privacy, and quantitative data methodologies including data linkage. He also leads the UK Anonymisation Network, which is planning to publish an anonymisation decision-making framework for those handling personal data and needing to make practical decisions about anonymising data, in May 2015.

2. Data Access Report

The Governance of Data Access report has been finalised (circulated to group prior to meeting). EAGDA approved it for sign off.

The group discussed dissemination and publication of the report. As well as publishing on the funders' websites and disseminating through research networks, the group felt that it would be important to publish a version of this report. The funders agreed with this approach. Options for seeking publication in *Nature* or *Science* were discussed, which were considered appropriate given that the report provides a practical counterpart to previous aspirational publications such as the Toronto statement.

The report would need to be edited to make it shorter and more appropriate for an international audience, rather than geared specifically towards the EAGDA funders.

EAGDA also asked the funders' whether they would be submitting a joint response to EAGDA in response to the report. Although each funder has different policies and would implement the recommendations in slightly different ways, it was agreed that a joint response would be valuable in providing EAGDA with key action points to monitor.

ACTION: Secretariat to explore options for approaching *Nature* and *Science* with the report with a view to publication via these channels.

Funders to discuss a joint response, to be sent to EAGDA in advance of the next meeting.

3. Nuffield Council on Bioethics report on uses of health data

A letter to EAGDA from the Nuffield Council, highlighting the recommendations from their recent report into uses of biological and health data of relevance to EAGDA, has been circulated to the group.

Several issues were raised for discussion, including: the recommendations around on-going participation on issues of ethics and data governance; the regulatory environment including the European Data protection Regulation; proportionality of governance mechanisms; the

role of consent; evidence of harms associated with data misuse; and costs associated with engagements and governance activities.

It was agreed that EAGDA should respond. The response will welcome the Council's engagement with these issues but also raise concerns about some of the recommendations particularly in relation to the role of consent in the revised Data Protection Regulation, differing interpretations of the risks of harms arising from the use of data and the need to consider the cost and resource implications of the requirements that the report advocates.

ACTION: Secretariat to draft a response to the Nuffield Council, with input sought from EAGDA members.

4. Work plan

Topics in the work plan were discussed:

Quantifying risk

Following the jointly commissioned evidence review for the Nuffield report, the group discussed what, if any, further work is needed on the issue of quantifying risks to data subjects of participating in research. The group concluded that it should be standard practice to estimate risks to research participants of different levels of data aggregations and for research studies to take this into consideration in developing their data sharing/release policies. However, risk assessment needs to take output conditions (i.e., the context and environment in which the data will be used) into account as well as input conditions (i.e., the data themselves), and so risk assessment can be challenging in practice.

The UK Anonymisation network is in the process of developing practical guidance on risk estimation informing data release and management. EAGDA will be await the publication of this guidance before deciding on any possible further actions on this subject.

ACTION: Mark Elliot to circulate UKAN anonymisation framework to the group when available (estimated May 2015); EAGDA to consider whether/how to take the issue of quantifying risk forward in the work plan at next meeting.

Legal control

A short briefing on issues related to legal control and responsibility for data access, linkage and storage (particularly with reference to cloud storage) in international contexts was circulated prior to the meeting. EAGDA agreed that this is a topic that merits further investigation and is of direct interest to the funders and the research community, especially in light of the growing usage of cloud storage. There is lack of uniformity with language and definitions used to describe the way data is controlled, with variations in different regions (e.g., the term 'data ownership' is unhelpful as it may imply data that one owns or is responsible for, or that refers to the fact that data refer to a particular person). This lack of clarity may lead to oversights or a failure to recognise issues – particularly where data are shared and/or stored across borders.

Scoping work will be required to hone precisely what the legal terminology, current controls and basic principles of legal control are across different data sharing jurisdictions.

Bill Lowrance was suggested as a possible author who would be well-placed to undertake this kind of research.

ACTION: Secretariat to draw up a specification for research into legal issues in relation to linked datasets and international cloud storage of/access to data and identify several possible appropriate researchers to approach.

Trustworthy research environments

The MRC has been taking forward the follow up to AMS's workshop last year on safe havens and is seeking to develop a set of parameters around how data should be handled, transmitted and used that would be applicable across different disciplines and settings. There has been confusion over what constitutes a 'safe haven' but a set of key principles and exemplars of good practice could promote research as trustworthy for handling and using data. The key point to note is that there are many different manifestations of what MRC terms a 'trustworthy research environment' that will be appropriate in different circumstances, and these may include virtual environments.

There are various strands of work currently being undertaken on this issue, including:

- ADRN is mapping convergences between different approaches to safe havens.
- Scotland is establishing five facilities that could be considered trustworthy research environments.
- The Global Alliance regulatory and ethics working group is undertaking similar work to identify and explore properties of safe havens and is developingprinciples that should be considered as criteria for safe havens in a genomics context.

The group discussed whether trustworthy research environments would be the most effective way of harmonising and streamlining access to data, and that they would need to be careful to avoid adding unnecessary bureaucracy to data access. It was also noted that training and accreditation would need to be included in any work to consider how to develop these environments, as the infrastructure has to be supported by well-trained scientists and data managers if they are to be effective.

ACTION: MRC to lead on developing a set of principles for trustworthy research environments, with EAGDA providing input as needed.

Consent

It was agreed by the group that consent was not an issue that EAGDA could usefully add value to, owing to the great deal of other work being conducted in this area. EAGDA will keep a watching brief on this issue.

Current legislation

Beyond biomedical data, a great deal of work is being done on legislation regarding access to information and the balance between privacy and public benefit, e.g., through surveillance and counter-terrorist measures. These measures may impact on public perceptions of research data sharing as they converge on the issue of privacy, which has not been discussed in the context of data reuse. The EU DPR in its current draft is still relying on the notion of personal data which is now outdated given advances in technology and the possibilities of re-identification.

It was agreed that EAGDA should keep a watching brief on legislation that concerns data access and reuse outside the domain of scientific research.

Sanctions
EAGDA has discussed the issue of enforcing sanctions both for data misuse (e.g., through deliberate attempts to re-identify individuals) and for failure to share data where appropriate. The funders noted that they had examined their respective policies on research integrity and agreed that it would not be possible to produce standard wording for sanctions across those policies.

It was agreed to hold the topic of sanctions for the next meeting to allow Bartha Knoppers – who proposed sanctions as a topic for EAGDA – to feed in views on whether and what further EAGDA could do in this area.

Commercial access to data
Andrew Morris raised interactions with the private sector on data sharing as a key issue that will only increase in importance over time. It is often perceived by research participants as a contentious area. There may be different models for engagement that may be more acceptable to the public and research participants. It was noted that some may see this as incompatible with the public sector ethos of the NHS and may lead to objections about the use of patient or participant data in commercial contexts. EAGDA agreed that these are important issues but there was not any key action point to take forward at this time.

The Wellcome Trust is currently commissioning social research to explore public attitudes to commercial organisations accessing health data; the ESRC is also looking to scope out where further work might be needed in this area; and MRC are participating in an initiative run by the Ministerial Industry Strategy Group about biomedical research interactions with industry.

ACTION: Funders to keep EAGDA informed about progress of current work in this area and identify potential issues that EAGDA could follow up on.

Infrastructure
Data sharing infrastructures need to be sustainably resourced over time, and need to have policies that are fit for purpose around e.g., retention and destruction of data. It was noted that the Software Sustainability Institute should be engaged on any funder discussions on this issue.

ACTION: ESRC/WT to lead on a workshop on the sustainability of data management systems and infrastructure, to report back to EAGDA at the next meeting.

HEFCE and data outputs in the REF
Several studies are currently mining the impact case studies from the REF 2014, but not many datasets were submitted in these. It will be important to understand the experiences of REF panel members in assessing datasets, and the perspectives of researchers and institutions, if they are to be considered as valued and valuable research outputs in the next round of research assessment. HEFCE is due to open a consultation for the community in the autumn on the next REF, and the importance of including datasets should be fed into this consultation.

EAGDA highlighted to the funders the need to push for engagement with HEFCE at this early stage.

ACTION: ESRC/WT to lead on a cross-funder roundtable with HEFCE on data outputs and the REF, and to report back to EAGDA at the next meeting.

International dimensions of data sharing
Mark Guyer reminded EAGDA of the need for UK funders to ensure that their policies on data-sharing and associated infrastructures are framed with the international context in mind. EAGDA must ensure all its deliberations have an international dimension.

ACTION: Secretariat to draft a letter from EAGDA to the funders on the importance of considering international aspects when developing policies and processes for data-sharing, This could either be included in a letter to the funders outlining EAGDA's proposed work plan or be included in future correspondence when EAGDA is corresponding with the funders on a relevant subject.

5. Update on Global Alliance

Paul Flicek provided an update on the activities of the Global Alliance. Policies on privacy and security, and on consent will be released later this month.
There is a plenary meeting in Leiden in June, which now has the backing of a large number of organisations beyond the US, Canada and UK, and is supported by enthusiastic volunteers. Sustainability will be the key issue in the long term and it is a positive step that a global network is developing around genomics. The group discussed whether the Global Alliance could be a model for other branches of science beyond genomics in future.

6. Updates from EAGDA members

Tim Hubbard provided an update on progress with Genomics England, which has received research ethics committee approval for its main programme. There are 11 genome medicine centres (consortia of participating hospitals) which are ready to go live and samples are now being collected. The data centre for feeding back findings into the clinic is also established,

and the Genomics England Clinical Interpretation Partnerships that will fulfil the research function of the programme are now being formed.

Melanie Wright provided an update on the Administrative Data Research Network, which has now gone live, linking government administrative data for research purposes. Many projects now have a biomedical element and so the ADRN is working with Farr Centres to ensure there is cross-disciplinary expertise to handle biomedical/biological data.

Mark Guyer indicated that there would be an update on NIH's Big Data to Knowledge and precisions medicine initiatives at the next EAGDA meeting.

Paul Flicek informed the group about progress with Elixir, an initiative to create a research infrastructure for life sciences in Europe. 15 countries have subscribed to the programme, which aims to co-ordinate activities across Europe and pilot/demonstration projects are underway.

7. Funder updates

Vanessa Cuthill outlined ESRC's activities since the last EAGDA meeting:

- A strategic plan published in January sets the emphasis on interdisciplinary working and collaboration with other funders.
- A new framework for research ethics for ESRC researchers is live, with case studies and exemplars being added.
- ESRC is funding a new centre for real-time data analytics
- Doctoral training provision is being redesigned to be more targeted and with better provision for training biosocial researchers.
- Work is being done to include social science research in discussions about the Data Protection Regulation.
- ESRC has been closely involved in the Cabinet Office open policy-making process on data sharing between and beyond government departments. The particular funder interest has been in the sharing of de-identified data for research and statistical purposes. The process is currently on hold because of the upcoming election.

Fiona Reddington outlined CRUK's recent activities:

- CRUK is in discussions with the EBI regarding the Big Data landscape and what role CRUK could play in facilitating this area – any potential proposals in this area would need to fit with other ongoing activities in this area such as the Global Alliance
- CRUK is working with Nature to develop clinical trial data descriptors, to assist with the discovery and sharing of clinical trial data.
- CRUK is currently investigating how to facilitate clinical trial data sharing amongst its funded researchers, and options are being investigated as to what support (including infrastructure) would be required to support this.

Jon Fistein outlined MRC's recent activities:

- MRC have been working on clinical trial data discoverability, and will stipulate in their policies that all trials should be registered, with summary results published whatever the outcome of the trial.
- Jointly with BBMRI, MRC are launching a tissue directory initiative
- MRC ethics guidance is currently under review.

Katherine Littler outlined the Wellcome Trust's recent activities:

- The Trust has taken over the review panel for clinicalstudydatarequest.com, the clinical trials data sharing platform set up by a consortium of pharmaceutical companies and led by GSK. It will shortly also take over the secretariat for the panel, which will be newly appointed.
- The Trust recently commissioned two reports in the area of data access and sharing:
    - A report by Technopolis on the research potential of clinical trial data;
    - A report in behalf of the Public Health Data Sharing Forum by a consortium of researchers led by the University of the West of England, on data linkage in public health and epidemiology.

  Links to both reports were circulated to EAGDA prior to the meeting.
- The Trust is in the process of commissioning social research to explore public attitudes in relation to commercial access to health, biomedical and genetic data. The samples will hopefully include a subset of cohort participants. There should be an interim progress report to provide to EAGDA in time for the next meeting.
- Update on discussions with regard to the DPR

8. AOB

The next meeting date for 2015 was confirmed as:

Wednesday 30 September 2015

Dates for 2016 will be confirmed before the next EAGDA meeting.

CLOSE