

## Better use of data in government consultation

### Responding to the consultation

#### Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional): Dr Natalie Banner

Position (optional): Policy Officer

Organisation name: Wellcome Trust. Also responding on behalf of MQ, British Heart Foundation and Cancer Research UK

Address: 215 Euston Road, London, NW1 2BE

Email: [n.banner@wellcome.ac.uk](mailto:n.banner@wellcome.ac.uk)

Telephone (optional): 020 7611 8235

#### Would you like us to treat your response as confidential?\*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes  No

#### Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

---

**If you ticked “Official response”, please respond accordingly:**

**Type of responding organisation\***

- Business
  - Charity
  - Local authority
  - Central government
  - Wider public sector (e.g. health bodies, schools and emergency services)
  - University or other higher education institution
  - Other representative or interest group (please answer the question below)
- 

**Type of representative group or interest group**

- Union
  - Employer or business representative group
  - Subject association or learned society
  - Equality organisation or group
  - School, college or teacher representative group
  - Other (please state below)
- 

**Nation\***

- England
- Wales
- Northern Ireland
- Scotland
- Other EU country: \_\_\_\_\_

( ) Non-EU country: \_\_\_\_\_

**How did you find out about this consultation?**

( ) Gov.uk website

( ) Internet search

(✓) Other

\_Open policy making process run by Involve\_\_\_\_\_

**May we contact you for further information?**

(✓) Yes ( ) No



## Cabinet Office: Better Use of Data

### Response by the Wellcome Trust, British Heart Foundation, Cancer Research UK and MQ

22 April 2016

#### Key Points

- We welcome the intention of the proposals to improve the efficiency, extent and transparency of data sharing within and beyond government. They need to be backed up by real incentives to promote an appropriate data sharing culture.
- We can only “unlock the power of data” and realise its benefits if the systems for managing and using data are trustworthy. This means that the processes for data access and sharing are transparent, with clear public accountability.
- Health and social care data should not become subject to further conflicting or fragmented regulatory frameworks. If they are beyond the scope of these proposals, it is imperative that these data are nonetheless able to be linked with other government data for research purposes.
- The proposals overlook existing infrastructures for securely and effectively handling and linking data, outside of government bodies.
- We welcome the inclusion of criminal sanctions for unlawful disclosure of personal data, but if they are to act as a deterrent whilst not stifling legitimate research uses, it needs to be made clear how and when they would apply.
- We are concerned that the definition of research in the public interest is very narrow and would restrict much research that is of value to society.

#### Introduction

1. We are major charitable funders of biomedical research in the UK. The various types of research that we support include cohort studies, which collect and link biomedical, health and other types of data from large numbers of individuals over time to enhance our understanding of health and disease. We believe that the responsible use and sharing of data is vitally important for research and the development of evidence based healthcare and services.

2. We recognise the enormous potential that improved data sharing from government departments has for the research that we and other funders support. The proposal to introduce legislation to simplify, standardise and promote the accessibility of government data is a welcome step, but one that needs to give serious consideration to the concerns outlined below if it is to succeed.
3. Our response includes remarks about the scope and principles set out in the proposals, followed by more specific responses to the consultation questions on **allowing the use of data for research and official statistics**.

### Scope and context

4. We recognise that health and social care data has always been beyond the scope of these proposals. However, paragraph 28 in the consultation implies that health and social care data may in fact be integrated into this legislation, with additional safeguards and protections on account of their higher degree of sensitivity. We urge the Cabinet Office to consult closely with the Department of Health and the National Data Guardian for Health and Social Care to integrate these proposals with legislation governing the use of health and social care data, and the upcoming Caldicott Review.
5. Lessons are still being learned in the health sector about how to build a trustworthy system for managing and sharing data, in which the public and data users can have confidence. We believe the Cabinet Office should draw on the experiences of the Health and Social Care Information Centre with the *care.data* fiasco and Partridge Review two years ago.
6. Not all data linkage will be within the confines of data held by public bodies. Much valuable research can be undertaken if data held by public bodies can be linked with publicly available data, or data gathered by researchers. For example, linkage between government administrative data and consented individual level survey data would allow for research that would enhance the richness and utility of both types of dataset. These types of linkage should be clearly stated as within the scope of the legislation.
7. The European General Data Protection Regulation (GDPR) is mentioned in the consultation, but without consideration as to what the impact of this Regulation on the proposals might be. Although we recognise that the details of implementation of the EU GDPR are well beyond the scope of the proposals to consider, we encourage the Cabinet Office and Department of Culture, Media and Sport to work together to promote a consistent approach between the proposed legislation and implementation of the GDPR.

### Building on good practice

8. The proposals do not adequately take into account existing infrastructures beyond government and their expertise in safely and securely managing data from multiple sources. The Administrative Data Research Network<sup>1</sup> (ADRN) and the Farr Institute<sup>2</sup> have both been set up to manage, link and enable access to different types of data. At present the ADRN is under-utilised precisely because government departments are not configured to enable the data they hold to be shared.
9. The processes for data linkage should not be specified in either primary or secondary legislation. The 'Trusted Third Party' is one model, but there is much innovation in this area, for example through using federated data access mechanisms. Requiring the procedures to be specified in

---

<sup>1</sup> [www.adrn.ac.uk](http://www.adrn.ac.uk)

<sup>2</sup> [www.farrinstitute.org](http://www.farrinstitute.org)

secondary legislation (paragraph 107) will create bottlenecks, stifle innovation, disincentivise the development of new methodologies and delay best practice developing as technologies evolve.

### **Transparency, accountability and trust**

10. If government data sharing is going to be trusted, it is critically important that the principle of transparency is embedded in the proposals. Transparency is a valuable safeguard, both to ensure that proper processes are being followed by all and to guard against the perception that data could be used for inappropriate purposes, either by government or by other organisations. Reporting mechanisms are needed for what data is used, by whom, for what purposes and what outcomes result from that data use. The culture of transparency needs to be fostered by all producers and users of data, even though it inevitably takes time, resource and a shift in culture to adhere to it.
11. Accountability also underpins trustworthiness. The consultation is unclear on the roles and responsibilities of different bodies in implementing the legislation in practice, including the Information Commissioner's Office, the UK Statistics Authority and the Office of National Statistics, and also other bodies beyond government that may hold and process data. If it is unclear who is accountable for what data sharing practices, all stakeholders will lose out: data producers and users will be unsure of their rights and responsibilities and the public will lose confidence in the system.
12. The Codes of Practice need to ensure the existing regulatory environment does not become more complex and fragmented. A map of information flows indicating who is the data controller and what responsibilities they have would help ensure that processes are transparent and there is full public accountability within the system. This will provide reassurance that government-held data is being used for public benefit and also provide clarity to potential users of data.
13. The private sector plays a role in delivering many public services and is also often part of the research process. Recent research about health data commissioned by the Wellcome Trust indicated that, when informed about the way data could be used, most people were open to the idea of commercial organisations accessing health data on the condition that there was a clear public benefit. There were also clear red lines about what people did not find acceptable, and a wide range of strongly held views.<sup>3</sup> Honesty and openness about the private sector using government data and what access is permitted or prohibited would greatly enhance the trustworthiness of the data sharing proposals.
14. We support the intention to create a system of accreditation for data users, data indexers and facilities. However, it is not proportionate to require individual research proposals for accessing de-identified data to go through an accreditation process if the individuals undertaking the research are themselves accredited and an appropriate, accredited access facility is used. If individual research proposals will in fact be required to go through a process of accreditation then substantial resource needs to be dedicated to this function. Otherwise, bottlenecks in the process will unnecessarily delay research. Lessons should be learned here from the experience of HSCIC's Data Access Request Service<sup>4</sup>.

---

<sup>3</sup> "The One Way Mirror: public attitudes to commercial access to health data" (Ipsos MORI/Wellcome Trust, March 2016): [www.wellcome.ac.uk/publicattitudes](http://www.wellcome.ac.uk/publicattitudes)

<sup>4</sup> The Wellcome Trust produced evidence for the Health Select Committee in March 2015 on the impact of delays at HSCIC on research: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/health-committee/handling-of-nhs-patient-data/written/18669.html>

## **Sanctions**

15. We welcome the proposal to introduce criminal sanctions for the unlawful disclosure of data, modelled on the provisions contained within existing legislation. Given the broad range of data covered by the proposals, a clear definition of 'unlawful disclosure' is needed that enables data producers and users to have confidence about what data sharing activities would be lawful and what would be unlawful. Without this clarity, it is likely that appropriate and legitimate data sharing will be inhibited rather than encouraged.
16. Whatever sanctions are introduced, health and social care data needs to be subject to the same safeguards and deterrents from misuse as other types of data.
17. Sanctions also need also to be proportionate and sensitive to the cause of the breach: different sanctions should be available against individuals deliberately misusing data, and organisations whose poor data security and governance lead to data breaches.
18. There is no mention of implementing any mechanisms for redress for individuals in the case of data being misused. This overlooks an important mechanism for ensuring accountability. We strongly urge the Cabinet Office to reconsider this oversight and to include measures for redress in draft legislation.
19. The proposals draw a clear dividing line between identified and de-identified data. However, the identifiability of data is a function of its environment: data that is de-identified in one situation may be identifiable in another if additional information can be linked to triangulate identities. Whether or not data is de-identified therefore depends in part on how it will be used and what other information it will be linked with. The proposals should take this context into account and not assume that data can be labelled as de-identified without consideration of the environment in which it will be used.

## **Resourcing and data quality**

20. Legislation alone is not enough to change embedded cultures and ensure government departments enable appropriate access to high quality data for research purposes. Infrastructure, expertise and data skills, incentives and resources are needed to support these activities. We would welcome further emphasis on developing the capacity and capability of the systems, organisations and people responsible for data sharing. There should be provision within the proposals to support departments to improve data quality and resource data sharing activities adequately.

## **Response to specific consultation questions**

### **Question fifteen**

21. In general, fees should not be charged for the service of providing data for publicly or charitably funded research purposes. For commercial users, it may be appropriate to implement some form of cost recovery system, with all fees for fulfilling the data access request quoted upfront. We take no view on the maximum fee that should be permitted, but consider that it should be proportionate to the costs of the data processing and other data service-related activities as required.

### **Question sixteen**

22. We agree that the UKSA should publish high-level details about rejected applications for access to data. The information made available should include the aggregate data on where requests are coming from, departments that data has been requested from, and the reasons for rejection. It

should not include academically sensitive information, for example about specific research questions asked, as this may put researchers at a competitive disadvantage.

23. Publication has several advantages. It can: highlight common problems or pitfalls with applications that other data users can then avoid; demonstrate to the public that inappropriate uses are not being permitted; and provide a benchmark, flagging any departments that are consistently refusing to allow access to data. The UKSA should compile this data in one easily accessible place to allow a snapshot view for the public and potential data users.

#### **Question seventeen**

24. We agree that “research in the public interest” should not be defined in legislation. However, the criteria for identifying research that has the potential for public benefit set out in the proposals are extremely narrow and would rule out a large swathe of valuable research. “Public interest” is not restricted to the improvement of public services or evaluation of public policy. Research using government data could, for example, involve increasing understanding of important social, economic, or environmental determinants of health.<sup>5</sup> Such research does not have a definite policy output, but is nonetheless strongly in the public interest.
25. We therefore advocate a much broader view of research in the public interest. Rather than being defined by the content of the research question (as indicated in paragraph 105), we suggest the Cabinet Office should explore alternative views. For example, public interest could be identified procedurally, as research approved as being “in the public interest” through a transparent process by a panel of peers and independent experts, the outputs of which are made openly available. This latter point would exclude research conducted solely for commercial purposes.
26. Alternatively, the Cabinet Office should consider setting out what would be excluded under its “public interest” definition rather than attempting to capture all of the nuances of research that would be permitted. The Wellcome Trust’s research on attitudes towards commercial access to data found that most people had a strong implicit sense that research conducted solely for profit-making was not in the public interest, but that commercial involvement in research that was seen as having a public benefit was generally accepted. Insurance and marketing uses of data were widely considered to fail this public interest test. If such data uses would be prohibited, these exclusion criteria could provide public assurance that research deemed as being in the public interest is acceptable and worthwhile.
27. It may also be helpful to consider how ‘public interest’ will be interpreted in implementation of the General Data Protection Regulation to promote a consistent approach.

#### **Final Remarks**

28. As medical research funding organisations, we believe that – although it is not the focus of these proposals – they are an important step in promoting the use and linking of government-held data to improve the nation’s health. We would be happy to discuss any of the points raised in this document further with the Cabinet Office.

---

<sup>5</sup> One example of this type of research comes from the Scottish Health and Ethnicity Linkage Study (SHELs): examining inequalities in all-mortality, all-hospitalisation, infectious and parasitic diseases, injury and poisoning, and bowel screening. This research links ethnicity and socio-demographic data from the 2001 Census in Scotland with health data from the Scottish Bowel Screening programme, in order to examine ethnic variation in bowel cancer screening in the Scottish population. The research has contributed enormously to understanding health inequalities and fulfils a strong public interest function.