# Invited response to the Nuffield Council on Bioethics report on ‘The collection, linking and use of data in biomedical research and healthcare: ethical issues’.

# Delivered at the Launch of the Nuffield Council on Bioethics Report 3 Feb 2015, Westminster, London.

Good afternoon. My name is Anna Middleton and I am a senior staff scientist from the Wellcome Trust Sanger Institute in Cambridge. My career has explored, from multiple different perspectives, the impact of genetic technologies on people. I have been asked to comment on this Nuffield council report because my current research has gathered the views of just under 7,000 people from 75 different countries on how they think genomic research studies *could* be conducted. I’m going to talk about this work, and explore how it fits with two of the themes identified in the report.

However, before I do that, I wanted to start by highlighting the *relevance* of Big Datasets specifically to the discipline of genomics. Genomics refers to the study of a person’s 20,000 or so genes. Given the *almost infinite* ways that people can be genetically different to each other, genomic research often needs to be done on a very large scale in order to be able to interpret the significance of, particularly a rare genetic change. So, Big Data and Genomics go hand in hand.

I’m part of the Deciphering Developmental Disorders or DDD project which seeks to offer cutting edge genomic testing to 12,000 children from the NHS who have an undiagnosed developmental disorder, i.e. a severe, complex, physical and/or intellectual disability. Such children have exceptionally rare conditions that their clinician may never have seen before. With the support of a particular online database called Decipher that contains large sets of health and biological data, the DDD project has managed to match specific genetic results with children at opposite ends of the UK, and has found a new diagnosis in more than 30% of the children tested so far– children who were not able to get this via the NHS. This would not have been possible, at the scale and rate we have seen so far, if there hadn’t been the prior collection, linking and storing of biomedical data from similarly affected children.

In the DDD project at the Sanger Institute we have access to large volumes of genomic data relating to thousands of patients. It goes without saying that this data is protected with the highest levels of security. Whilst the project aims specifically to understand the genetic architecture behind developmental disorders we also have access, should we so choose to look, at all sorts of other incidental health data about each individual participant, e.g. we could explore genes related to cancer, heart disease or alzheimers. Whilst we don’t store personally identifiable data on campus and the researchers are not able to identify individual patients themselves, if something of significance was identified it could be shared with a patient’s clinician via the use of a unique code and the clinician could contact the patient. In the DDD project we have taken the ethical stance that only data relating to the developmental disorder will be returned to patients. The social science project that I want to share with you explores what potential research participants expect of researchers with regards to receiving personal genomic data.

The 7,000 or so participants in our study were members of the public, health professionals and genomic researchers. Some knew absolutely nothing about genomics before seeing the survey, others were experts in the field. A particular hook that brought people to our survey was that it contained 10 short films within it that describe the ethical issues raised by genomics, these sit within the survey and create an opportunity for participants to think, learn and reflect. The films helped the survey to go viral online, which is how we were able to reach people from across the world.

We asked participants to imagine that they were a research participant in a genome sequencing study. We explained what this was, we explained the sort of data that a genome study looks at. Did they expect to receive personal results? If so, what sort of data were they interested in? What would they do with it? Did they expect researchers to keep analysing their data? Did they want a flexible or dynamic consent process that enabled them to update the information they received or didn’t receive over time?

Whilst all social sciences data should be reproduced in multiple populations before drawing any broad conclusions, our survey offers a first attempt, on a large scale, to explore attitudes in this area. In our data analysis we worked very hard to minimise bias and adjust for all potential confounding effects.

Irrespective of where people lived in the world and we have participants from 75 different countries, including Russia, UK, the US, France, Australia, India, China, S.Africa, the overwhelming message was that participants were excited by the prospect of genomic research. They wanted to participate, they wanted to contribute to the global understanding of human health and they welcomed the opportunity to learn new information about themselves.

98% of participants said that if it were possible to receive genetic information linked to serious or life-threatening conditions that could be prevented – they would want to know. They were interested in data for personal use to protect their health. 94% were interested in knowing information about how they respond to medications or drugs, 43% were even interested in uncertain information that cannot be interpreted at the moment. There was a sense that ‘if you know it, I’d like to know it too’.

Even if participants didn’t understand what it meant or how it would be useful, they wanted to be given the opportunity to learn and be connected to the research outcomes. What we don’t know is whether they just happened to be a very positive group of people and would have been as positive about research in other areas of health – but I can’t imagine the same excitement about getting access to an X-ray of a knee. We explained clearly the limitations of genomic data and yet still given this caveat, participants cared that genomics research happened. Clearly this has relevance for future management of expectations.

When asked to explore their reasoning, participants were hopeful that genomics would make a difference to health. However, they also said, in very high numbers, that were concerned about the potential for insurance companies to use genomic data to discriminate. Yet, this concern was not enough to stop them wanting genomic research to proceed. A word about insurance: in many countries across the world now individuals are protected from discrimination by insurers on the basis of genetic test results.

Despite the perceived negative outcomes of genomic research there was still a strong indication, both through the quantitative data gathered, but also via the free text comments, that participants hoped to be involved in research design. They didn’t want to be a detached partner – simply providing a sample and then forgotten about. They wanted to be a respected associate in the process, part of the conversation. This reflects well with sentiments provided in the Nuffield council report. I’m going to highlight two specific quotes from the Reflections and Conclusions section. The first is about the role of ‘participation’ in research. It reads: “Participation demonstrates respect for persons by involving them in the design of data initiatives (it enables them to engage in forming the conditions of a future in which they have a direct interest rather than merely responding to it) and is more likely to produce outcomes that secure their commitment and build trust.” This chimes with our study results. The second quote comes at the end.

What was really interesting is that people were able to put a value on how important the data was to them and the majority said that if the delivery of personalised data meant that this compromised the researcher’s ability to conduct their research and answer their research questions then they did not expect data to be given to them. This speaks of a respect for the researcher and research process – participants were not belligerent in requiring or insisting researchers delivered data to them, but were conscious there might be consequences attached to their requests. They appreciated and supported the necessity of researchers to do good quality research and did not wish this to be compromised.

Research participants were supportive of researchers interrogating their data as much as possible – they wanted it to be useful to the research exercise. They were also supportive of being able to engage in a flexible or dynamic consent process where they could alter the levels of information they could access.

So, in conclusion, the common themes throughout our data, irrespective of geography, age or demographic background was that participants were supportive of genomic research and they wanted to be part of the journey to inform human health.

I feel our findings are reflected in the sentiments of the Nuffield council report and I’d like to end with a quote, again from the Reflections and Conclusions section, this really speaks to the data that I have collected:

“The key to acting ethically with personal health information in a world of Big Data will be to maintain the engagement of, and oversight by, patients and other affected people not just as a new initiative is being developed, but as it evolves over time”.

I would like to conclude by offering my wholehearted support to the council’s recommendation that social sciences research on both the benefits and harms of research for individuals, groups and society are conducted, to further inform this debate. Ethical and moral reasoning cannot exist in a vacuum; it has to be connected to the practical reality of what is feasible and the social reality of what this means to people.

We need to continue to ask people, in a multitude of ways, how they want to engage with the issues highlighted in the report.