Debating Data: A Citizens' Jury

February 15th 2020, Oxford Town Hall

Background

Who ran the event?

This public engagement event was run by a team from the <u>Wellcome Centre for Ethics and Humanities</u> (WEH), based at the Big Data Institute, Oxford and the <u>Centre for Personalised Medicine</u> (CPM), a partnership between St Anne's College and the Wellcome Centre for Human Genetics, Oxford. Both the WEH and the CPM are interested in people's views of the uses of health data in research and the provision of clinical care. WEH researchers undertake research on ethical issues arising in the clinic and are involved in a number of biomedical research projects, which involve the sharing of MRI images, digital pathology images and genomic data. The CPM is involved in public engagement work in personalised medicine and involves a number of people working within the field of genomics and its application in personalised medicine. Given their overlapping interests, members of the two centres came together to collaborate on this public engagement event on health data sharing. The organising team included the Public Engagement Manager from the WEH (Milly Farrell), members of the WEH (Nina Hallowell), CPM (Padraig Dixon, Jason Torres) and two 'Industrial Strategy Challenge Fund (ISCF) Centres of Excellence' research groups; the National Consortium of Intelligent Medical Imaging (<u>NCIMI</u>) (Danielle Wilson) and the Northern Pathology Imaging Cooperative (<u>NPIC</u>) (Daljeet Bansal).

Why data sharing?

Given the speed that data-intensive technologies are advancing it is important to regularly gauge publics' views of the applications of technology in healthcare. Considering the WEH's and CPM's interest in the use of health data in research, we thought it important to determine publics' views of the sharing of different types of health data (MRI, digital pathology images and genomic data). Specifically, we were interested in views on the benefits and challenges of sharing these data for research purposes, the acceptability of sharing health data with commercial companies and other ethical issues raised by data-sharing.

Aims

The **Debating Data** event had two related aims. The primary aim was to pilot a deliberative public engagement method – a citizens' jury. Neither the WEH nor the CPM had run a citizens' jury before, and both were interested to trial this method of determining publics' views about complex healthcare issues. The secondary aim was to gauge adult (18+) publics' views on data-sharing, specifically comparing and contrasting the sharing of different types of health data with public and commercial entities.

What is a Citizens' Jury?

Citizens' Juries usually involve a representative sample of citizens (n= 12-24) who are brought together to discuss an issue, or set of issues, over a number of days (2-7days). Within the jury format, citizens are briefed with evidence from expert "witnesses", whom they

interrogate. Citizens then deliberate over a set of predetermined questions. Finally, they are asked to come to a consensus position, or produce a "verdict". This form of engagement has a number of strengths and weaknesses. The strengths include: the enabling of direct public input into decision-making/public policy, the ability to interrogate experts and evidence, and the support for discussion of pre-specified outcomes. Weaknesses of the method include small sample size (although this can be offset by representative sampling) and the fact that they involve predetermined issues of interest to policy makers and researchers rather than those which may be of greatest interest and relevance to the citizens themselves[1].

Recruitment

We used various means to recruit our participants. Our intention was to recruit an adult cohort (+18 years) of mixed demographics to encourage a range of viewpoints. We advertised in the local Oxfordshire 'Daily Info' newsletter (both print and online) and advertising flyers were circulated at local hospitals and other healthcare waiting rooms in the region. Closer to the event date, we approached University colleagues facilitating PPI events to advertise the event to members of their existing databases. Attendance was free and all catering and travel costs covered. Registration numbers (via Eventbrite) were capped at 35 to allow for potential drop-out on the day.

Timetable

Participants arrived from 09:30 and final verdicts closed at 16:30. Participants collected name badges with an allocated number and colour to denote respective deliberation groups. A member of the planning team facilitated group deliberation during Sessions 1-3. The day comprised of four sessions:

Session 1: 'Types of Medical Data'- Three expert speakers (Simon Leedham, Fergus Gleeson and Richard Colling) gave short presentations on the basics of the data types (Genomics, Medical Imaging and Pathology Image data) relevant to their respective research/clinical work. Participants then broke-out into their 'numbered' groups and discussed the following questions:
- In which ways are these different types of data similar and why?
- In which ways are they different and why?
- What relevance do these types of data have for you?

---

[1] Information taken from INVOLVE citizen-jury methods webpage:
https://www.involve.org.uk/resources/methods/citizens-jury (3/6/2020)

Session 2: 'Sharing of Data'- The three speakers gave short presentations on the basics of how their respective data set is currently used and shared. Participants then broke out into their colour groups and discussed three pre-written case studies about the use of different types of data (see 'Materials').

Session 3: 'Use of Data and Ethics'- One presentation was given by Bioethicist (Michael Parker). Participants then broke out into their 'numbered' groups once again and discussed the following questions on the ethics of data sharing:
- What ways should these different types of data be used?
- Should these different types of data have different 'rules' about how they're used and shared?
- Who should/should not be able to access to your data?

Session 4: 'Verdicts'- While remaining in their 'numbered' groups, the participants were asked to reflect on a set of 4 key questions and try to reach a consensus (see 'verdicts'). No facilitator was present at this session and the participant groups worked independently to reach their verdicts. Each group nominated a speaker who then presented their verdict in the Council Chamber to all participants and facilitators. These verdicts were voice recorded for use as evidence in this report.

Outcomes

The Jurors

Prior to the event 35 individuals registered to attend. 20 attended on the day; 8 men 12 women. The age of the participants ranged between 18–70+ years. Twelve participants were

residents in central Oxford and 4 resided elsewhere in Oxfordshire, 6 came from outside of the region.

Although we advertised the event in different media across Oxford, the individuals that took part were not a representative sample of the population of Oxford. We realised our recruitment strategy – print, social media and word of mouth – would attract a self-selected sample. This prediction was borne out by the fact that facilitators and presenters recognised a number of participants as frequent attendees at public engagement events in the Oxford area and three participants indicated that they had heard about the event through the University of Oxford PPI network.

Main themes of Sessions 1-3 deliberations

The following provides a brief synopsis of the main themes arising in the different discussion groups. In Session 1 the groups discussed the meaning of the data types and their similarities and differences.

Similarities across data types.

All types of data were considered personal and as raising issues of data security. Participants felt that individuals have little control over health data. It was noted that all three types of data are used in determining diagnosis and treatment and are dependent upon technology.

Differences between data types.

Genomic data were regarded as different from MRI and Pathology image data because they are: predictive, shared with family members and potentially identify individuals. Image data were seen as fixed in space and time and personal in that they only pertain to individuals rather than biologically related groups. These different qualities mean that genomic data should be treated differently to the other data types.

In the remaining Sessions 2 and 3, a number of themes recurred across the groups in response to direct questions about data-sharing and when discussing the case studies.

Control and consent

Who decides how data is used? Participants discussed control of health data and how to establish individuals' consent for its use (i.e. should it be opt-in or opt-out), how should it be obtained and who could and could not consent and, in the latter case, what can be done going forward. Questions were raised about maintaining control of data uses if commercial entities are using the data. It was suggested that publics should be involved in debate about designing consent procedures for data access and should have a place on data access committees

Data ownership and benefit

Who owns the data and who benefits from its use, the individual or society? Participants regarded data as valuable assets. It was clear that participants did not feel that ownership of health data should be ceded to commercial entities, particularly big tech companies as, to quote one participant, "big companies scare us". It was observed that the NHS should retain ownership of data and a share of the IP generated by the data and should recoup costs by charging for data access and other benefits from data sharing with commercial entities, in particular.

Privacy and security

Issues of confidentiality, data privacy, data security and oversight of data uses were discussed as was the quantification of risk associated with different data uses and the need for trust and transparency regarding who is making decisions about data uses.

Public understanding

Finally, it was noted by our participants that wider publics may have little understanding of AI and, consequentially, asking technical questions about data sharing and access for the purposes of creating AI systems may be unproductive and ineffectual given the knowledge gap.

The verdicts

The participants were allocated to one of four groups at random, and asked to reach a verdict on the basis of the evidence given and the deliberations they had engaged in during the day. They were asked to respond to the following 4 questions when considering their verdict:

1. How should genomic data be used, who should they be used by and why?

2. How should imaging data be used, who should they be used by and why?

3. How should pathology data be used, who should they be used by and why?

4. Can you identify any key differences between the three types of data that you would like to mention?

They were asked to reach a consensus opinion if possible, although dissenting opinions were recorded. All groups reported a consensus was reached.

There were a number of overlapping themes across the 4 groups' verdicts, namely:

- **Genomic data is different** from MRI and Pathology image data and deserves greater protection
- Patients **should** give consent for use of their health data
- **Commercial uses** of health data should be allowed but **should be subject to more stringent oversight** than public uses

Evaluation

'Very insightful, enjoyable and met some great people' - Participant

The majority of participants were positive about their participation in this citizens' jury, with the exception of a lone detractor, who questioned the motivation for the event, suggesting

that it had been organised in order to persuade the public to agree to data sharing. Most (80%) said the event was enjoyable, liked the venue (70%), felt the event was well organised and the format of the day worked well (70%). Seven respondents commented on potential organisational improvements:

"Wider cross-section of participants, and more of them. More open-ended discussion. Groups to discuss issues raised by the presentations and then return to question presenters." - Participant

"Having breakout sessions with 'syndicates' can be a bit hit and miss, owing to the personalities in each group, but this is inevitable. Having a facilitator for each group during the individual sessions (not just the final one) would be very helpful in minimizing the tendency to go off the topic." - Participant

Nine respondents felt they had learnt something about health data and felt better informed following their participation:

"Learned much more about underlying science; felt better able to take an informed view" - Participant

"I learned a lot about the different types of data and how they are used" - Participant

A number of respondents (n=9) felt that the aim of the day, to understand public' views about the health data-sharing, had been met:

"I think this event absolutely met this aim, and I thought that the idea of having us all reach a conclusion at the end, which we shared in the final session, was brilliant" - Participant

However, some of this group felt that although the aim of the day was met, there was too much information to take in and/or too little time for discussion:

"Broadly, yes. However, although the speakers were excellent, there was too much information to take in. Perhaps more case studies would have helped." - Participant

Finally, one of the respondents said that they felt as though the information they received was biased and there was not enough focus upon the wider politics of health and data use more generally:

"[T]he event failed to contextualize data gathering and uses. The public have [sic] insufficient information about the uses and sharing of their data. To limit the discussion to NHS research purposes is to avoid the bigger issues of the politics of health. The event attempted to simplify the debate by reducing it to three types of data use and sharing without fully recognizing the wider issues." - Participant

When asked to reflect overall on their participation in the citizens'' jury, most of those who responded (n=7) were positive:

"This was the first time I've participated in an event like this. I enjoyed this first experience, and met some nice people. I'd be interested in doing more." - Participant

However, one participant was less than satisfied commenting that they:

"Feel manipulated and expected to endorse health data use and sharing after hearing the benefits of research in those three fields. The issue of who has access to our data and for what purposes cannot be confined to health research. If the responses from the day stifle a wider and informed public discussion on uses and sharing of health data, without acknowledging the controversial and political implications of the issue, then it will have been a cynical exercise to promote health research at the expense of democracy. The Data Guardian has recommended that the public are [sic] better informed. This event has not satisfied that recommendation." - Participant

Despite the fact that all of the materials used were framed in a neutral way and presenters were asked to provide a neutral account of data and data sharing in their presentations, when taking this forward we will stipulate the need for the witnesses to take a neutral stance, as in this instance, and check all materials, including presentations, as well as providing jury members with more time to interrogate witnesses, in an effort to overcome perceived bias.

Conclusions

Because the primary aim of this event was to pilot the citizens' jury method, we were working within the constraints of a limited budget (i.e. ~10% of the normal costs of running a citizens' jury[2]). This affected the outcome in two ways. First, it meant that we did not attempt to convene a representative sample of the population of Oxford. Second, it meant that we could only afford to convene the jury for a single day, rather than the recommended 2-7 days. In combination these adaptations mean that our results are unrepresentative and preliminary, in that the participants did not have enough time to thoroughly consider the questions we asked. However, given these constraints the results nevertheless provide some insight into what a self-selected sample of citizens think about different uses of different types of health data.

First, the participants argued that citizens should be able to control uses of their health data and give consent for data use in research. Second, participants regarded genomic data as more personal than image data and therefore, as requiring greater protection and more stringent oversight. Finally, participants believed that commercial researchers should be able to access health data but should be more closely monitored than publicly funded researchers.

The evaluation suggests that in the main, participants were very positive about taking part in this event. The organising team also felt that the event had gone smoothly, had delivered some interesting pilot data on publics' views of data-sharing and most importantly given them experience of how to run such events in the future.

With regards to lessons learnt for the future, it is important to: engage with a representative sample, allow more time for evidence, questioning, group work and deliberation and focus on straightforward questions. On reflection after the event, we felt we had set too ambitious a goal for one day by asking participants to compare and contrast three data types and deliberate on their appropriate usage by both public and private researchers. If the jury had taken place over several days, or perhaps three weekends it is possible that this multipart

---

[2] Information taken from INVOLVE citizen-jury methods webpage:
https://www.involve.org.uk/resources/methods/citizens-jury (3/6/2020)

question would have not seemed so complex as participants would have had more time and opportunities to grapple with these issues.

While the verdicts and discussion may only give us a preliminary view of publics' views on healthcare data-sharing, the organising team were pleased with the outcome and are planning to use this method in future events, albeit adapting recruitment methods and providing further training for facilitators beforehand. The CPM may use this format to engage with the public on aspects of data-sharing for precision medicine in London and Oxford. The NPIC team in Leeds are planning citizens' juries on the sharing of digitised pathology images for AI development in West Yorkshire and the WEH is planning at least one event, potentially with a younger public cohort to explore differences in generational perspectives on data sharing. The issues that were raised by the public and findings from this report may also inform further research on publics' views of data sharing. This method can therefore facilitate a cohesive 'listening loop', where future research on data use and sharing can be co-produced with relevant public participants, ensuring the research is fully transparent, inclusive and of benefit to all stakeholders.