**Access to biobanks: responsibilities within a research ecosystem**

**Abstract**

The last few decades have seen rapid increases in the size and scope of biobanks, with large-scale, publicly funded ventures supporting health-related research becoming the norm. As these biobanks are increasingly asked to share their data, including for example, genome-wide analyses, questions arise about how such decisions are made, including whether applicants’ research aligns with the aims of the biobank. To better understand how biobanks make decisions relating to their data use, we sought the views and experiences of those involved in decision-making relating to data access at 11 large-scale publicly funded health biobanks. We were particularly interested in how potentially contentious applications were approached. Interviewees had some concerns about decisions on applications they felt their governance structures could not reach. We ask broader questions about the responsibility of those involved in biobank access decisions –actors working more broadly early in any research process - when considering such issues.

**Introduction**

The last few decades have seen rapid increases in the size and scope of biobanks, with large-scale, publicly funded ventures that support health-related research becoming the norm. Many of these biobanks were established to provide a repository of population/participant/patient biosamples, and their associated data, for access by a wide range of research institutions and/or commercial research ventures.[[1]](#footnote-1) Advances in data collection, analysis, storage, and processing have allowed these biobanks to move from being primarily focused on acting as a [bio]sample repository, to sharing vast quantities of data from the analyses of these samples. For example, new genome sequencing technologies mean that data from whole exome and genome analyses (and other ‘omic analyses) are available for many samples. Whilst access to samples was always to be tightly regulated given their finite nature (repeat samples could be requested but may not provide the same information if temporally distant), sharing of data from these samples can be done repeatedly without reducing their utility to others, so could be perceived as less problematic. However, questions about access remain. For example, opportunities to link genomic data with other data such as clinical or social data mean that promises about anonymity made at recruitment may be difficult to fulfil. Moreover, biobanks are often set up with a particular aim in mind - for example, to further health-related research - so those governing access need to ensure this is what the researchers aim to do, yet this can often be difficult to delineate from, for example, if social, behavioural or ancestral implications emerge from health-related research.

The Ethics Advisory Committee of the UK’s largest publicly funded health-based biobank (UK Biobank), and one of the largest publicly funded health biobanks in the world, wanted to better understand how biobanks make decisions about applications seeking access to data, especially those that are potentially contentious, for example, those that are deemed to sit at the boundary of health-related research. Our research question was: what are the views and experiences of those involved in decision-making regarding data access at large-scale, publicly-funded biobanks that support health-related research?Health related research could be a primary or other aim**.** Privately funded biobanks were not considered for analysis because it was assumed that their decision-making processes might be influenced by different business models. We conducted interviews with 15 individuals associated with data access decision-making at 11 different biobanks.

All interviewees reported that only a minority of applications for data access were potentially contentious, and they said that their biobank had governance structures in place to handle such applications. Despite this, some interviewees narrated their concerns about making decisions regarding applications that their governance structures could not reach. Specifically, interviewees explained that while they could clearly reject some applications, others were only potentially contentious. Governance structures provided processes to deal with these applications but could not prevent health related research arriving at contentious, non-health-related conclusions. Similarly, they could not prevent media inflammation or distortion of research findings. We discuss what responsibilities can be placed on those working in the necessarily early, or upstream process of data access governance, to ensure there is no eventual inappropriate dissemination of research.Before we present our findings, we provide a brief overview of Data Access Committees – a key aspect of biobank access procedures.

**Data Access Committees**

A transparent and accountable procedure for access to a biobank’s resource is vital (1-4). Data Access Committees (DACs) have been established to monitor access and authorise requests and are viewed as an important element of such procedures (1, 5, 6). However, there are no internationally recognised standards for the role of DACs[[2]](#footnote-2) (5) and studies have shown signiﬁcant variation in the type, scope and structure/location of DAC committees. Some biobanks have scientiﬁc review boards or management committees that scrutinise access applications (1, 7).

Nevertheless, the literature notes some common criteria that DACs apply to submitted research applications. These generally (though not always)[[3]](#footnote-3) include ensuring that the data and/or samples are only shared with *bona ﬁde* researchers, and that the researchers need to be afﬁliated with reputable institutions that can be held responsible for their actions. The planned research must also meet the terms of a biobank and applications are usually rejected if they do not (2, 5). Research applications that are deemed objectionable have beendescribed in a range of ways, summarised as research that ‘*may impact or harm dignity of the human beings in a way that is undesirable or unacceptable in a democratic society’*. Examples of objectionable research stated include culturally or politically sensitive topics; ancestry studies in a small, isolated population; correlating cognitive ability and education to race or sexual orientation; genetic studies of addiction and mental health; and brain size correlations with certain social identities or geographic locations (5, 8). Below, we go on to show how, in practice, making decisions about whether access applications fall into the remit of objectionable research can be tricky, with some interviewees expressing concerns about how to make appropriate decisions in these circumstances.

**Methods**

*Identification of large-scale publicly funded biobanks*

We found no accessible list of large-scale biobanks. Therefore, and, following other studies (9), we searched a range of databases to identify suitable biobanks: the dbGAP[[4]](#footnote-4) website was searched for datasets/biobanks that comprised over 10,000 participants; and the BBMRI[[5]](#footnote-5) website was searched for population-based biobanks for general health applications, with these biobanks then being manually checked for those databases over 10,000 participants (this was because there was no search algorithm for the size of a biobank, and it was considered that population biobanks would be some of the largest biobanks in the BBMRI). In addition, general online searches and snowballing were conducted, including biobanks referred to in the literature and during interviews. For example, an article was identified on the web that highlighted some of the largest biobanks in the world (or on their way to becoming so), and so we checked this and other similar articles for publicly funded biobanks to determine if they would be relevant for inclusion.[[6]](#footnote-6)

Only English language webpages were looked at and only those with accessible contact information. Nineteen biobanks were emailed. These included biobanks from continental America (n=7), Europe including the UK (n=9), continental Asia (n=2), and continental Australasia (n=1). Emails were sent to e.g. the Chair of the access committee, a general biobank email address, or all members of a biobank access committee (total emails: n=30). After two follow up emails to non-responders, 15 individuals agreed to participate from 11 different biobanks.[[7]](#footnote-7) We are aware that this strategy will have missed a number of large biobanks not visible through websites in English. This is a limitation of the study, but we were not aiming to research an exhaustive list. Instead, as is typical of qualitative research, we aimed to explore the individual/group decision-making process about potentially contentious research , rather than to conduct a quantitative survey of different biobank governance structures. The aim was to generate new insights that could then be contextualised in the literature.

*Interviews*

Interviews were conducted virtually, digitally recorded, and lasted between 37 mins and 90 mins. Interviews sat at the boundary between semi-structured and unstructured: the interviewer used an interview schedule but was also led by interviewees if the discussion moved elsewhere. The interviewer returned to the interview schedule at various time points to ensure all main questions had been covered by the end of the interview. Interview questions included asking interviewees about the details of their biobank and the type of samples and data it collects, stores and processes; interviewee’s role/expertise within the biobank and also, where relevant, outside of the biobank; and their views on the benefits and risks of biobanking. Interviews also explored how decisions about access are made. This included questions about the membership of their access committee; governance structures; how and when meetings take place; how committee members know and interact with each other; and the value of the different types of committee members. It also included questions about who reviews access proposals and why, how a proposal is reviewed (which factors and/or values are considered and how are they given meaning during the review process) and views about the usefulness of criteria used during review (e.g. ‘health-related’, ‘public benefit’ etc). Finally, interviewees were asked about their experiences of handling potentially contentious applications (for example, those that sat on the boundary of what they would consider as health research), what these applications looked like (with examples requested), how discussion around them was held, and how any disagreements, uncertainties or concerns were discussed.

Interviewees were part of a biobank’s internal access team (n=4), part of the biobank data access committee (n=5), part of both (n=5), or had an overarching role – including the consideration of access - at a biobank (n=1). Interviewees were from biobanks in six countries and four continents (Europe, Asia, Australasia, North America). Interviewees’ disciplines were either biomedical research (n=3), health informatics/data analysis (n=2) or epidemiology (n=1); or bioethics (n=1), policy-related (n=2), or managerial (n=3). Three interviewees were participant representatives.

*Analysis*

Analysis of interview data was conducted by GS. Analysis was inductive and completed in two inter-linked rounds: broad coding, followed by detailed coding of the transcripts (11). Specifically, initially, interview transcripts were carefully read and re-read for relevant ideas and themes, and combined with the extensive memo-making taken directly after interviews. Second, interview transcripts were analysed line-by-line using NVivo software. Coding was carried out via constant comparison, which was continual, rigorous and allowed for developing themes. Coding and themes were discussed in detail with AL, throughout the analysis process.

The disciplines that interviewees came from is not compared in the findings as this made no differences to their responses. Interviewee identification is withheld if, for example, they refer to a specific practice or process that was absent in [most] other biobanks.

*Ethics Approval*

This study received ethics approval from the University of Southampton, the Faculty of Medicine Ethics Committee: ERGO 62731.R4.

**Findings**

All interviewees described how their biobanks streamline their access procedures to facilitate access, and aimed to encourage and support researchers to use the resource’s data. Although this facilitation was a key role, some interviewees explained how there were certain types of research/researchers that their biobank would not support even if it fell under the broad umbrella of their biobank’s remit. For example, while nearly all interviewees were associated with biobanks who permitted industry to access their data, all stressed that access would not be permitted for ‘*a purely commercial endeavour’* (interviewee 12) or an endeavour that involved data being sold to third parties (interviewee 4) (some described receiving such applications). One interviewee noted how research that supported interests of the tobacco industry or the promotion of alcohol was also prohibited in their biobank (interviewee 1). Non-acceptable research topics included for example, social research that involved biobank data being linked to social data such as school grades, jail sentences, salaries and education attainment data. Two interviewees described rejecting such research because it raised questions about the relationship between genetics and social standing, attainment and behaviour (interviewees 8/4). Interviewee 11 spoke about the data access committee discussing proposals from researchers who hold politically unacceptable views:

there are questions sometimes about the person, when they look them up and find that they’ve published views that are politically unacceptable, and so they may have extreme political views that are racist or sexist and have been public with those.

Furthermore, two applicants explained how their committee had deemed those applying for access who came from institutions that seemed to be associated with non-health purposes (*‘places that look like they’re military operations in [country withheld]’* (biobank interviewee 11)) not supportable, and another two interviewees explained how commercial enterprises that had little information on their website to be able to judge the credibility of the research had also been flagged as problematic.

Applicants also described research applications that they found difficult to make decisions about whether to grant access. For example, in the extract below, interviewee 5 described a research application, in which the applicant wanted to explore aspects of air pollution. While air pollution was viewed by the access team as health-related, the interviewee explained that there had been discussion around whether the findings of this research could have led to negative ramifications, and what responsibilities the access committee might have in relation to this. For example, if the findings of the research pointed to a specific location or area that has a low amount of air pollution, this could lead to the building of infrastructure in that location/area, which could drive up air pollution. This worried this interviewee:

somebody was looking at stuff to do with air pollution but we thought that that actually would have been health-related…It wasn’t, it’s not directly health-related but it was in a way, in that the hypothesis being that if you’re living in a more polluted area then this may have an effect on your health..[..].., the air pollution data may actually be used to justify the building of a new motorway, because ‘it’s fine, because this area is not going to cause too much pollution because there isn’t any’.

In another example, interviewee 13 reflected on an application that was related to mental health:

substance abuse or mental health disorders in [a stigmatised community], the way we have been reporting has always been very discriminatory. It has been published like, therefore, [this community] are drunks, they have a high disposition to be alcoholics, and things like that.

However, interviewees explained how such applications were rarely seen: ‘*for now, we didn’t face a tricky case..[..]..as I mentioned, it’s a straightforward procedure for us’* (interviewee 2); ‘*so far we haven’t had anything like this [research at the boundary of health-related research]...no, no such example at all*’ (interviewee 4).

**Handling ‘tricky’ applications: governance structures**

Nearly all interviewees reported that the access procedure for their biobank was a two-tier system. An internal access management team would typically review applications to ‘*screen the project’* (interviewee 2). Once applications had been reviewed, depending on the biobank, the access committee would view these proposals.[[8]](#footnote-8)Committee membership (between two and nine members) was diverse and often included a member of the internal access team and one or more biobank management personnel. Larger biobank committees tended to meet regularly, smaller committees worked by soliciting the opinions of two or more individuals that had been invited to act as committee members. When committee members needed to consider potentially contentious applications, interviewees described regulations and governance structures to assist with decision-making. For all interviewees, discussion between committee members was vital; ‘*we have a lot of discussions about that’* (biobank interviewee 8); ‘*we do discuss things, and we do discuss things in quite a lot of depth, actually, and everyone has to agree’* (biobank interviewee 6). Discussions involved considering participants who had provided their data to the biobank: ‘*would…participants…be surprised or be pleased or be alarmed by any of these proposals?*’ (biobank interviewee 14); *‘the access committee should be the voice of the people that have made the donation’* (biobank interviewee 15).

Participants described how these discussions could sometimes be resolved by asking applicants for more information about the potential value of the research proposals. As interviewee 3 explained, *‘we’d have to know why people [researcher applicants] wanted to do what they were going to do’.* This interviewee continued: ‘*[applicants need to] bolt on some sort of description of what [they] might do with it [the research] that’s helpful*..*[because]..[..]..we [the DAC] really need to think about what we’re [as a society] going to do with it’* (biobank interviewee 3). When this was not the case, other mechanisms were in place. Two interviewees opined that, when faced with a potentially contentious proposal, it was important for the biobank to seek broader input from a wider community by consulting the views of those community groups who were most likely to be impacted by the research. Interviewee 13 describes the mechanism for doing this within their own biobank. They explained:

we have these experts that - sometimes they are participants themselves or the community itself…that is how much we include, or we are sensitive to potential misuses or contentious uses of the resource. We hope….through the process, [the views of] those who are being impacted by the decision [are represented].[[9]](#footnote-9)

For (other) interviewees, if discussion failed to resolve the issue, then a governance structure (or governance structures coming into place; biobank interviewee 9) to allow for formal escalation of discussion and/or decision-making was drawn upon:

in case there is a dispute in the access committee decision [in the future; there had not been one yet], it can be referred to the regulatory committee. The regulatory committee refers its decision to the board.…So we have this hierarchy (biobank interviewee 2).

In one instance, for example, an interviewee explained that ultimately the final decision-maker would be the biobank regulator:

eventually the one who makes the decision is the biobank director…. if we are [still] not sure about something we can also ask for guidance from the authority that governs biobanks. We have one case where we ask them because we are not sure (interviewee withheld).

**Concerns about reputational damage**

Several biobank interviewees linked their concerns about potentially contentious applications with the need to safeguard their biobank’s reputation: ‘*it’s…also safeguarding the reputation of [the] biobank’* (biobank interviewee 7). Some interviewees (though not all, e.g., biobank interviewee 15) explained how permitting access to their dataset for a potentially contentious application could lead to a situation in which the news media picked up on the research using an inappropriate sensationalist framing:

*you would…have to address what happens when it all goes wrong, on Twitter and in the [UK] Daily Mail [newspaper]. You are allowing…access to all our genomic data*… (biobank interviewee 3).

As biobank interviewee 6 emphasised, *‘we’re always very concerned that we don’t attach ourselves to a company or a proposal that’s going to put [our biobank] in a bad light’.* Not being seen in a bad light was viewed by all interviewees as linked to public trust in biobanks and biobanking, and this was something that was perceived by nearly all interviewees to be vital to the biobanking endeavour; *‘just to get this trust from public…You have to be pretty sure that people believe what you are doing and you are doing it correctly’* (biobank interviewee 4)(12).

Interviewees emphasised the importance of their governance structures that had been put in place to handle these types of issues: *‘don’t start collecting data within six months…spend two years on getting your systems right, and your follow-up right’* (biobank interviewee 12). These were viewed as vital for accountability, and the justification of decision-making, especially in case a decision was later questioned:

*I try to work on you don’t legislate for the worst-case scenario. You legislate for the majority, but you have processes in place, and you have the documentation, and you can, at least, go back and say, ‘Well, this was the decision we made, and here’s the documentation around it’. I guess, if you get it wrong, sometimes you have to put your hands up and say, ‘Actually, that was the wrong decision’* (biobank interviewee 10);

*we would want to be able to justify not only to the participant panel but to the cohort in general, why we do what we do, and if there was a particularly controversial thing that came up and we made a decision one way or another, we would want to be able to justify that* (biobank interviewee 5).

One intervieweedescribed how their biobank had developed a specific approach in an attempt to address the issue of reputational damage. This interviewee explained how their biobank access committee ‘*encourag[ed] or nudg[ed] researchers to have a good public dissemination plan in the proposal’*:

*there should be further dissemination plans, could you specify, or some of the risk, of the social risk, that the group harms could be mitigated by having dissemination practices that go back to the community or pay attention to how your findings will be disclosed or could be published in the scientific territory but also the layperson territory. The headlines, the newspaper...[..]…we try to educate them. It’s not perfect…[..]..just to address the potential for headlines such as…’that research project said that poor people live in the more toxic environments and therefore, they will get cancer and risk in terms of being discriminated against’. Things like that, or ‘they will also have the intellectual disabilities, the children, because they live in toxic places and therefore…’ - those kinds of connections* (interviewee withheld).

However, there was a realisation that governance structures could only go so far, and that this alone would unlikely protect a biobank’s reputation because *‘it’s a tricky problem…it’s a judgement call...there isn’t a really firm decision that you can make about it, and you can’t go off a checklist that perfectly captures [it]’* (interviewee 10). This was because, as interviewee 3 remarked, *‘you never know what’s contentious until someone starts contending it’*. Interviewee 7 concurred: ‘*what threatens reputation is very difficult to foresee*’.

Overall, these interviewees spoke about the importance of having transparent and accountable governance processes they could point to in the event that the reputation of their biobank was called into disrepute. Nevertheless, there was consensus that these governance processes could only go so far, and issues associated with predicting what research would be contentious at such an early stage of the research process were raised. We discuss these further in the below.

**Discussion**

Our findings suggest that the representatives from large-scale, health related, publicly funded biobanks we spoke to see relatively few applications that they deem potentially contentious. Regulations and governance structures helped interviewees to handle any that were. However, some interviewees remained concerned that despite these, certain decisions remained difficult. For example, those that had a potential to lead to societal harms, not only to biobank participants and/or specific communities, but also to the biobank itself in terms of reputational damage. There is a need to understand how to proceed responsibly when faced with difficult decisions about these potential societal impacts of research.

Not all the thinking related to the potential societal implications of research can be done at such an early stage of the research process. Indeed, concerns have been raised elsewhere that those with a gatekeeping function so early in the research process struggle to enact responsible innovation practices as they relate to societal impacts. For example, Samuel and colleagues (2021) have shown that within the Higher Education Institution (HEI) research ethics ecosystem for population health artificial intelligence (AI) research, there is a separation between the ethical practice of research, and research use (societal impact), and that the purview of HEI research ethics committees focuses on issues of research ethics, leaving little space and less oversight for considerations related to the ethical use of research. Here, research ethics is seen as synonymous with data governance, and considered procedural; societal ethical issues were seen as less tangible, including notions of societal harms that could come from the misuse of AI algorithms in practice. These authors argue that to address this, we must view responsibility as an aspect of a research ecosystem, rather than giving all responsibilities to research ethics committees and researchers (13, 14). In this ecosystem, all actors - whether biobanks or researchers (or funding bodies, research institutions, publishing houses, professional bodies etc) – should have a shared understanding of what responsible innovation looks like, and work to a shared goal. In this case, the authors viewed this shared goal as relating to the on-going scrutiny of research and its applications to ensure the mitigation of societal harms. These authors have therefore called for ex-post review of population health-related AI research to synergise thinking around the ethics of research and its impact (15).

Similarly, biobank data access committees, which are situated early in the research process, have a finite role in being able to mitigate potential societal harms, particularly around issues associated with news media discourses and reputational damage that are likely to be temporally distant. However, there are practices that biobanks can integrate into their governance processes to mitigate these harms as much as possible. First, they can adopt an ex-post review process for those proposals that the access committee have flagged as particularly potentially contentious so that they can learn about the research findings, and then work with researchers to ensure these findings are disseminated to the (scientific and) public domain in appropriate ways. For example, while researchers have little control over the news media’s framing of research once research is in the public domain, one biobank attempted to counter this by educating the researchers they support about news media relations at the stage of application, and this is a practice that could be adopted by other biobanks, perhaps at an ex-post review stage.[[10]](#footnote-10) Furthermore, some biobanks had participants on their data access committees, which are an important means to ensure such views are represented through participatory governance. We recommend participants’ views are consulted through such committee or subcommittee membership for all large-scale biobanks.[[11]](#footnote-11) While individuals cannot be expected to represent the diverse interests of all biobank participants, they can provide important insights into a wide range of participant perspectives. If participant views about potentially contentious proposals differ to those of other committee members, consensus must be sought and participant views should be given equal consideration to other committee members. Some of our interviewees described such a process in their data access committees. Interestingly, we were told how participant representatives often brought questions to the access committee as to why biobanks were not sharing *more*.

It is also important to realise that those working in news media also have a responsibility and should be held accountable just as much as those working in biobanks and other aspects of the research process, to ensure responsible dissemination of research findings that arise from a biobank’s data resource. News media professionals (public relations officers, journalists, editors etc) are just as much actors in the research process/ecosystem and just as much required to work in line with the spirit of responsible innovation.[[12]](#footnote-12) In fact, media scholars have long called for news media professionals to take account of the possible personal and social impacts of the representation of scientific research, arguing that such reporting needs ‘*to exhibit the virtues of humility and responsibility no less than clinical research and care practices do’* (16).

In summary, while biobanks have established governance processes to support access decisions, some of our interviewees worried about the societal harm that could result from their approval of access requests for a potentially contentious proposal. We emphasise that because biobanks represent an early-stage actor in the research process, they cannot govern and control later stages of the process. However, there are a range of processes that biobanks can put in place to help mitigate potential issues. These include consulting participants about proposals deemed potentially contentious and having an ex-post review procedure for such proposals. At the same time, we argue that responsibilities associated with research endeavours need to expand to include actors such as new media professionals, and need to be part of the ongoing research ecosystem over time.

Our findings are particularly important as biobanks continue to link genomic and clinical data, and the realm of what is considered health data is expanded and becomes ever more blurry (17). Biobanks of the future may well start ‘banking’ other health data, such as self-tracking data such as data from wearables, administrative data, environmental data (e.g., pollution sensor data), transactional data (e.g. supermarket data). In fact, such proposals for data collection are already underway.[[13]](#footnote-13) Understanding the enactment of responsibilities within such large repositories is the next challenge for biobanking.

**References**

1. Langhof H, Kahrass H, Illig T, Jahns R, Strech D. Current practices for access, compensation, and prioritization in biobanks. Results from an interview study. European Journal of Human Genetics. 2018;26(11):1572-81.

2. UK Expert Advisory Group on Data Access. Governance of data access. 2015.

3. UK Biobank. ACCESS PROCEDURES: Application and review procedures for access to the UK Biobank Resource. 2011.

4. Cadigan RJ, Juengst E, Davis A, Henderson G. Underutilization of specimens in biobanks: an ethical as well as a practical concern? Genet Med. 2014;16(10):738-40.

5. Shabani M, Dove ES, Murtagh M, Knoppers BM, Borry P. Oversight of Genomic Data Sharing: What Roles for Ethics and Data Access Committees? Biopreserv Biobank. 2017;15(5):469-74.

6. Gibson SG, Axler RE, Lemmens T. Transparency of Biobank Access in Canada: An Assessment of Industry Access and the Availability of Information on Access Policies and Resulting Research. Journal of Empirical Research on Human Research Ethics. 2017;12(5):310-25.

7. Kaye J, Briceño Moraia L, Mitchell C, Bell J, Bovenberg JA, Tassé AM, et al. Access Governance for Biobanks: The Case of the BioSHaRE-EU Cohorts. Biopreserv Biobank. 2016;14(3):201-6.

8. Langhof H, Kahrass H, Sievers S, Strech D. Access policies in biobank research: what criteria do they include and how publicly available are they? A cross-sectional study. European Journal of Human Genetics. 2017;25(3):293-300.

9. Shabani M, Borry P. “You want the right amount of oversight”: interviews with data access committee members and experts on genomic data access. Genetics in Medicine. 2016;18(9):892-7.

10. Tryka KA, Hao L, Sturcke A, Jin Y, Wang ZY, Ziyabari L, et al. NCBI’s Database of Genotypes and Phenotypes: dbGaP. Nucleic Acids Research. 2014;42(D1):D975-D9.

11. Strauss AL. Qualitative Analysis for Social Scientists. Cambridge: Cambridge University Press; 1987.

12. Samuel, G.; Broekstra, R.; Gille, F.; Lucassen, A. Public trust and trustworthiness in biobanking: the need for more reflexivity. *Biopreservation and Biobanking*. [10.1089/bio.2021.0109](https://doi.org/10.1089/bio.2021.0109)

13. Samuel G, Derrick GE, van Leeuwen T. The Ethics Ecosystem: Personal Ethics, Network Governance and Regulating Actors Governing the Use of Social Media Research Data. Minerva. 2019;57(3):317-43.

14. Samuel G, Chubb J, Derrick G. Boundaries Between Research Ethics and Ethical Research Use in Artificial Intelligence Health Research. Journal of Empirical Research on Human Research Ethics. 2021;16(3):325-37.

15. Samuel G, Derrick GE. Defining ethical standards for the application of digital tools to population health research. WHO bulletin. 2020.

16. Montgomery J, al. e. Novel neurotechnologies: intervening in the brain. Nuffeld Council on Bioethics; 2013.

17. Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. Bull World Health Organ. 2018;96(1):66-8.

1. The types of biosamples and data vary depending on the biobank. Biosamples may include blood, cheek swabs, urine and/or saliva, or for biobanks focussed on particular conditions – for example cancer - surplus tissue from medical investigations such as biopsies. Data might include health measurements taken at the time of recruitment to the biobank (blood pressure, height etc), or information from linked medical records, such as General practice or hospital records, or information from questionnaires (e.g. lifestyle, family history, environmental exposure, etc). [↑](#footnote-ref-1)
2. Though both the BBMRI and Global Health Alliance have documents under consultation that explores this issue (*personal communication*). [↑](#footnote-ref-2)
3. For example, the United States All of Us program has been established to be inclusive. Data can be accessed by researchers beyond those attached to a specific/reputable research institution. Nonetheless, researchers must still register to be validated, and therefore become ‘qualified researchers’. Furthermore, the biobank operates a tier based system, such that ‘qualified researchers’ may only access certain tiered data. [↑](#footnote-ref-3)
4. A National Institutes of Health-sponsored repository charged to archive, curate and distribute information produced by studies investigating the interaction of genotype and phenotype, https://www.ncbi.nlm.nih.gov/gap/ 10. Tryka KA, Hao L, Sturcke A, Jin Y, Wang ZY, Ziyabari L, et al. NCBI’s Database of Genotypes and Phenotypes: dbGaP. Nucleic Acids Research. 2014;42(D1):D975-D9. [↑](#footnote-ref-4)
5. BBMRI is a European research infrastructure for biobanking, which contains a searchable database of European biobanks. https://www.bbmri-eric.eu/ [↑](#footnote-ref-5)
6. https://www.biobanking.com/10-largest-biobanks-in-the-world/ [↑](#footnote-ref-6)
7. More than one individual replied to our invitation from various biobanks. [↑](#footnote-ref-7)
8. One biobank did not have an access committee, but would open up the proposal to wider discussion if necessary. In some biobanks, if the proposals were viewed as uncomplicated they were just ‘desktop’ reviews. [↑](#footnote-ref-8)
9. We have not provided additional information on this mechanism because it contains identifying information. [↑](#footnote-ref-9)
10. If resources permit, this could be offered to all researchers accessing biobank data and/or summary statistics. [↑](#footnote-ref-10)
11. If resources are an issue, participants can be consulted only for potentially contentious applications. Interviewees told us these applications were seen very rarely. [↑](#footnote-ref-11)
12. Social media now has a large role in news dissemination, and consideration of how this might impact societal discourses is also required. [↑](#footnote-ref-12)
13. See the UK social data foundation proposal at Southampton University websciences where there is an agreement between council/ university/ hospital to gather such data [↑](#footnote-ref-13)