Social and Behavioural Research in Clinical Genetics

Section Editor:
Aad Tibben, email: a.tibben@lumc.nl

What is a biobank? Differing definitions among biobank stakeholders


While there is widespread agreement on the broad aspects of what constitutes a biobank, there is much disagreement regarding the precise definition. This research aimed to describe and analyze the definitions of the term biobank offered by various stakeholders in biobanking. Interviews were conducted with 36 biobanking stakeholders with international experience currently working in Switzerland. The results show that, in addition to the core concepts of biological samples and linked data, the planned use of samples (including sharing) is held to be a key criterion. It also emerges that some researchers avoid the term in order to circumvent certain regulatory guidelines, including informed consent requirements. Developments in the field of biobanking will be complicated if researchers are unaware, or deny that their collection is a biobank. A clear definition of the term is therefore an important step towards fostering collaboration amongst researchers, enabling them to more easily identify potential sources of samples.

Conflict of interest
No conflicts of interest to declare.

Biobanks are now highly regulated in most countries, and many official national and international guidelines offer definitions of biobanks. These definitions differ in their details, but the majority state that biobanks are repositories of biological samples with accompanying linked data. For instance, the Organisation for Economic Co-operation and Development (OECD) defines biobanks as ‘structured resources that can be used for the purpose of genetic research and which include: (i) human biological materials and/or information generated from the analysis of the same; and (ii) extensive associated information’. (1) The UK Biobank Ethics and Governance Council has stated that ‘the most robust contemporary definition of ‘biobanks’ is rich collections of data plus biospecimens, specifically developed as resources for research’ (2). As in both these definitions, the term ‘biobank’ is normally used for repositories that are intended for use in research, rather than for diagnostic purposes. Using different terminology for different types of repository allows different regulatory regimes and consent procedures to be applied.

However, the extent to which biobank stakeholders are aware of and agree with official definitions is unclear. Given the many differences in biobank type, size, and research focus, and the lack of an accepted formal definition, it is unsurprising that numerous definitions also appear in the literature (3–8).

Definitions are an important issue because the laws and regulations designed to govern biobanks are unlikely to be effective if biobankers are unaware that their collections of samples are biobanks, or if they are reluctant to define them as such due to the regulatory requirements which can ensue. Furthermore, uniting researchers in collaborative networks will also be difficult if awareness about the existence, extent, and value of sample collections remains limited. In order
to investigate this issue, and other potential barriers to biobanking, qualitative interviews were conducted with 36 stakeholders with international experience currently working in Switzerland.

Methods

Interviews

In order to explore their definitions of a biobank, we conducted interviews with key stakeholders involved in biobanking. On the basis of a review of existing debate reported in the literature we developed a guide for conducting semi-structured interviews. We used purposive sampling, with the goal of identifying key stakeholders in biobanking in Switzerland through publications, biobank and academic networks, as well as personal contacts and a snowball approach. Among the stakeholders identified were biobank managers, pathologists, researchers, clinicians, lawyers and ethicists. Questions about definitions were included within a broader interview guide and were preceded by more general questions about biobanking activities. Our specific aim was to allow stakeholders to freely express their own definitions in response to the question ‘what is your definition of a biobank?’, and explore their criteria with follow-up questions. Interviews were carried out in person or over the phone based on the preference of interviewees and travel convenience, and were transcribed verbatim. Confidentiality fully guaranteed and transcriptions were anonymized in order to prevent identification through names or recognizable situations. A total of 36 individuals working in connection with biobanks in Switzerland agreed to participate.

Analysis

The transcribed interviews were read in full by four members of the research team (including all the authors). Data were analyzed following classical qualitative methods (9). Transcripts were coded to identify key themes, concepts, phrases and attitudes. The authors compared codes and agreed upon the organization of sub-codes.

Results

Biological material/data

The majority of participants share the opinion that a biobank must contain both biological samples and data; see Box 1 for some typical responses. Several participants said that biobanks must be ‘systematic’ or ‘structured’, where structured means ‘a reproducible way and a transparent way of how the things got in there’ (I27). Some interviewees made a distinction between a ‘repository’ or ‘collection’, which do not (need to) have data, and a biobank, which does:

A repository is… the difference between a repository and a biobank is, a repository is a collection of biopsecimens. Maybe also some data. A biobank… by definition have to have both: biospecimens and data. Otherwise it’s not a biobank. (I11)

In contrast, a minority of participants stated that a biobank need not contain data: ‘It’s an organization to store… biological samples for later usage’. (I19) In terms of the biological material itself, a minority of participants mentioned that the biological material need not be human: ‘…it’s… I would say it’s extremely broad… it’s any biological material being animal, plant… that is… You know located somewhere’. (I16) A few participants claimed that the samples must be of several different types of (human) tissue:

I would maybe restrict it to… to biological material… tissue and, and fluid… And I don’t know if I would go so far to say that every DNA or RNA or protein piece is a biobank. (I25)

Size

Most participants did not mention size as being part of the definition of a biobank, but a minority did think that size was relevant. Several participants stated explicitly that size could determine whether a group of samples constituted a collection or a biobank: ‘…it’s not a biobank to have ten samples in a freezer and Excel sheet. That’s a collection’. (I31).

Another participant said that whether a collection of samples was a biobank depended on how rare the disease in question was:

It very much depends on… I mean if you have a rare disease, even a few samples are quite valuable. If it’s a more common disease you may need several samples because, depending on… I mean, depending on the scientific question, you may not answer a question with ten patients. In hepatitis C for examples, it would be very difficult, unless you are looking at a very rare phenotype. (I17)
Use of samples

A substantial minority of participants mentioned the intended use of samples as being a component of the definition of a biobank. Among those who did mention the purpose for which samples would be used, all mentioned research (see Box 2 for examples). One participant expressed an objection to biobanks with no specific purpose:

I resent the idea of biobank without questions, ... I think that biobanks should be focused on questions. On scientific and clinical questions. (I32)

Box 2
Biobank stakeholders’ views on the use of biobank samples

I28: “And, let’s hope it’s used for research, or for ... clinical work.”
I24: “... a biobank ... is used to do research that has nothing to with the direct treatment of the patient anymore.”
I12: “... to provide researchers or clinicians with, with tissue, and data.”
I13: “... A biobank, this is, this is for me a place where you collect and store biospecimens ... that you can use at some point for starting some, some new projects.”
I36: “I just want to make this possible, research on biological samples, and this is how I would do it. That’s more the ... like the banker, not ... the banker he doesn’t have, he doesn’t own the money, he basically offers services, if people have money, he can handle that... and... make more out of it.”

The same participant stated that s/he had to ‘... fight with my hospital which is creating a biobank which I think is useless...’ (I32). Finally, the importance of diagnostic biobanks and of potential rediagnosis was also mentioned by several subjects.

Sharing

Some interviewees mentioned sharing as a key component of a biobank: ‘I think a biobank should be something that is used by different researchers, and collaborators also, it’s not something you are only doing for yourself’. (I30). Another participant thought that sharing was relevant to determining whether his/her collection actually was a biobank:

So, if you, that’s the question, if your definition of a biobank is any storage of samples, then yes, we have a biobank, if biobank means it’s so to say like external ... we have samples for anybody ... Then no, we have no biobank. (I26)

Avoidance of the term ‘biobank’

Several participants said that some biobank stakeholders (on some occasions including themselves) choose not to call their sample collections ‘biobanks’ in order to avoid regulation ‘because of the legal aspects’ (I22):

if I would call it biobank, I would get into a lot of troubles...because all these crazy regulations concerning with biobanking...The problem with biobanking is...these people who have brought up this, this topic, had thought, thought of, had invented... a whole, very long list of regulations, how you have to deal with biobank specimens, how the patient has to get involved yeah, and this for us, only meant increase of bureaucracy, complicated, complicated the whole thing... for example exchange of archival material, even anonymous, between departments, which was... which was a custom, which was a regular habit, yeah, is now very difficult because what is missing for our archive material is the... a written consent of the patient. (I22)

The same participant went on to explain why these regulations meant that very valuable samples might go unused:

...but we have all these, we have all these wonderful specimens with very rare tumours, and so, and these data, these, these... data security issues... now almost prohibits that we use them for quality control. That’s allowed, yeah?...but that’s... that becomes legally difficult. (I22)

Another participant also mentioned avoiding the biobank ‘tag’ because of the issue of consent:

people will start thinking about whether they want to be biobank or not...because biobanking then is... you know, it’s, to be a biobank, considerable effort.... I mean if you call it a biobank you would have to have informed consent forms. (I27)

Discussion

Biological material and data

The views of most participants were in line with the official definition of the Swiss Academy of Medical Sciences (SAMS, see below), although only one participant mentioned this particular definition specifically. As such, biological material and data were seen as the baseline criteria for a biobank. Interestingly, however, several interviewees suggested that some collections that contained both material and data might nonetheless not be a biobank, in contrast, some interviewees stated that even collections without data were technically biobanks; this may be because some pathologists’ collections are composed only of samples, with the data being stored elsewhere; it could be argued that such a collection only becomes a biobank when hospital data are linked to the samples.

Size

Most participants agreed that size was not an important part of the definition of a biobank, but a few claimed that a relatively small number of samples and linked data is simply a collection. This is in accordance with the UK Biobank Ethics and Governance Council criterion of a ‘rich collection’, but it is unclear why a ‘poor’ collection with linked data should not also be regarded as a biobank and subject to regulation. Furthermore, as another interviewee pointed out, the
value of a sample is linked to the rarity of a disease, meaning that size alone does not determine the richness of a sample.

Use of samples
Almost all participants who mentioned the intended use of samples stated that research is the main purpose of biobanks, which is in line with the UK and OECD definitions. Interestingly, one participant claimed that biobanks focused on a central question are more valuable than those used for general purposes, but unfortunately did not provide any clear reasons for this belief.

Sharing
Although sharing is not mentioned in most official definitions of biobanks, it is encouraging that several participants saw sharing samples with external institutions as a key aspect of a biobank, as this practice is widely promoted as a means of facilitating research (10–12).

Avoidance of the term ‘biobank’
Perhaps the most interesting finding is that several participants admitted that they would not describe their collections as ‘biobanks’ despite the fact that they did meet a number of the commonly held criteria. One noted that calling his collection a biobank would lead to him getting into trouble. To this interviewee, the perceived overregulation of biobanks means that adopting the term to describe sample collections brings with it a number of undesirable consequences. While choosing to avoid the term ‘biobank’ is not in itself problematic, carrying out research upon samples without ethics committee approval could be regarded as research misconduct. This participant and another admitted describing research as quality control in order to circumvent biobanking regulation. This is a common problem in research review generally: researchers wish to avoid the perceived complications of obtaining approval from an ethics committee, so instead attempt to label their study as audit or evaluation (13, 14). While this might be understandable from a pragmatic perspective, conducting research without the necessary approval contravenes the basic principles of the Declaration of Helsinki, and is against the law in many jurisdictions (as one participant seemed to realize) (15). The attitude of the interviewees in question appeared to be that this is just how things are done in their line of work; the lack of a standard definition allows such practices to continue, as the choice of what to call one’s sample collection is, to some degree, decided on a case-by-case basis. While a globally accepted definition of the term ‘biobank’ cannot in itself encourage individuals to call their collections biobanks, it may facilitate holding medical practitioners accountable for the use they make of certain samples.

Several participants also mentioned the burden of obtaining informed consent from patients. Of course, some sample collections are genuinely not biobanks and it would not be helpful or necessary to require informed consent for use of such samples. One participant expressed his frustration that regulations make it difficult to use valuable samples that were obtained without consent. A new law governing human subjects research in Switzerland comes into force on the 1 January 2014 and will allow the use of historic samples without patients’ consent, which could address researchers’ concerns to some extent (16). However, this option does require approval from an ethics committee, which means that there might still be some motivation to claim that research is ‘quality control’.

The importance of a clear definition
Our results make it clear that there is widespread disagreement about the definition of a biobank. This may have important implications for sharing and cooperation between biobanks and regulation of the use of biological samples. First, if a researcher thinks that his collection is not a biobank, he is unlikely to respond to any local, national or international communications concerning biobanks or biobank networks. Such researchers are therefore less likely to share samples with other biobanks or ask for samples to be shared with them. Second, if a researcher does not think he has a biobank, he will not think that any relevant biobank standards or regulation apply to him (see previous section). This means that the use of samples might not be regulated properly. While poor sample quality and inadequate regulation are also important barriers to sharing, ignorance or denial that a collection is a biobank can act as a barrier to sharing even if quality and governance are of a high standard.

What, then, is the best definition of a biobank? Remarkably, no previous paper has focused on the issue of exactly what a biobank is. As mentioned above, a wide range of definitions are provided by official bodies and in the literature. The results from our study confirm that samples and data are seen as the basic requirements for a biobank, with use of samples also identifies as an important aspect. Our interviewees generally felt that the size of the collection was unimportant, which reflects the majority of national laws and guidelines (17–19). If we are to arrive at an ideal definition, it will be necessary to decide whether size and/or use are essential components of a biobank.

The SAMS provides the following definition: “Biobanks are systematic collections of samples of human body substances (e.g. organs, tissue, blood, cells etc.) and DNA as carrier of genetic information. Data that contain information on the donor (demographic data, type of disease etc., but also genetic data) are stored, either together with the samples or separately” (20).

In fact, we would argue that this definition is actually superior to that offered by the OECD, because it
does not dictate that the information associated with biobanks must be ‘extensive’. A biobank can be a very valuable resource without having particularly detailed associated data or (as one participant mentioned) thousands of samples, and imposing such conditions would allow very many collections to remain outwith biobank regulation despite their importance for medicine. However, while it avoids this problem, the SAMS definition has the contrasting disadvantage of casting the net too widely: it would include any collection of routinely collected biological specimens with linked data, even if such a collection was never intended to be used for research.

As such, we would suggest that the best definition of a biobank would be one that does not refer to the size of sample collections or the richness of data, but does state that the purpose of the biobank is research. Adding the research criterion to the definition has the dual benefits of reflecting current biobank practices, and regulating what needs to be regulated without accidentally catching collections that would not require for their intended purpose. This ensures that any samples that are to be used for research will be regulated, but those used for quality control, diagnostics or forensics (none of which require consent) will not. This would also avoid the problem of researchers having to deal with regulation when they have no intention of using historical samples. Therefore, we tentatively suggest a definition such as the following: ‘A biobank is any collection of human biological samples and linked data that is to be used for research’.

A clear and widely accepted definition of biobanks would remove the ambiguity currently surrounding some collections, making it easier for researchers to determine whether any given collection is in fact a biobank. However, in addition to publicizing the SAMS or an alternative consensus definition, it may be necessary for individual institutions to conduct internal audits of collections to check that no biobanks are being overlooked; some institutions have established committees for this latter purpose (21). Such audits could also help increase the number of biobank samples available for sharing.

**Conclusion**

The results of this study reveal that there is substantial disagreement among biobankers about exactly what a biobank is. There is a general consensus regarding the key criteria of biological samples, data, and their use for research, but quite divergent views on the importance of size, sharing, and diversity of samples. Defining biobanks is an important part of regulating biomedical research, but the possibility that such definitions might be unwelcome for a variety of reasons has not been previously explored. Clear definitions of biobanks and education of biobank stakeholders may be necessary to facilitate future biobank development and sharing. Tighter regulation of non-biobank collections and sanctions for those who attempt to ignore guidelines may be necessary to prevent researchers exploiting this loophole.

**Acknowledgements**

Thanks to Heidi Carmen Howard, who assisted in designing the interviews, gathering the data, and the initial coding process.

F.C.’s research is supported in part by the Käthe-Zingg Schwichtenberg Fund of the Swiss Academy of Medical Science. All other support for all authors came from the University of Basel. This paper was not commissioned by, revised, or otherwise connected to this funding. The authors declare no institutional or commercial affiliations which could constitute a conflict of interest.

**References**