

Claudia Mund

## **Biobanks – Data Sources without Limits?**

*The combination of health and lifestyle data with bodily substances and genetic information has led over the last few years to the creation of so-called biobanks. These biobanks are used to research a large number of diseases of modern civilization and genetic interactions. The interest of researchers in a preferably unrestricted right to data usage clashes with the personality rights of data subjects. What restrictions should be imposed on data use, and how can this be done without rendering research activity impossible? Balancing interests – considerations from a data protection perspective.*

### **Table of contents**

- I. Introduction
- II. The issue of consent
- III. Anonymization and encoding of donor data and samples
- IV. Data access by third parties
- V. Building trust through a research confidentiality policy
- VI. Conclusions

## **I. Introduction**

[Rz 1] Type the word «biobank» into an Internet search engine today and the responses returned will include a large number of major projects involving research into the most common diseases of civilization of our time. One of the best-known projects is without a doubt the Icelandic biobank, operated since 1998 by a private pharmaceutical company, deCode Genetics, and set up with government support<sup>1</sup>. The database contains family histories and medical records, as well as biological samples taken from a large section of Iceland's homogeneous population. The hope is that this data will allow a correlation to be established between genetic predispositions and the onset of widespread diseases. Other examples of population-based biobanks include the Estonian Genome Project, launched in 2001 by the Estonian government<sup>2</sup>, as well as the UK Biobank which is due to begin work at the beginning of 2006 and will contain data and samples from approximately 500,000 donors<sup>3</sup>. A further biobank, PopGen, is set up in the state of Schleswig-Holstein in Germany<sup>4</sup>.

[Rz 2] What these data collections have in common is the combination of health and lifestyle data with human bodily substances and genetic information for the purpose of researching complex diseases of civilization. While such data collections are of great scientific interest for research purposes, they also are attracting a lot of interest from third parties. For example, the Swedish police investigating the murder of Foreign Minister Anna Lindh succeeded in obtaining access to a neonatal database in order to be able to compare crime scene traces with the blood samples of a suspect. The same database was used for the identification of Swedish tsunami victims with genetic information collected in the form of DNA profiles.

[Rz 3] This paper focuses on some of the fundamental problems connected with the use and possible combination of data from these so-called biobanks which contain health and lifestyle data, as well as human tissues and genetic information about donors. As it will be shown, there is a permanent tug of war between the interests of researchers, the rights of data subjects and the interests of third parties.

## **II. The issue of consent**

[Rz 4] Modern research relies on a wide range of personal data and on access to human bodily substances. Both, that is the personal health, lifestyle or genetic data, and the biological samples of the data subject, enjoy protection under the right to informational self-determination as well as the relevant national and international data protection provisions<sup>5</sup>. As these personal data are sensitive ones, they can only be processed with the explicit and informed

consent of the data subject concerned. The «informed consent» requirement could, however, constitute a problem if a pre-existing data collection is being used<sup>6</sup>, or if a research project moves in a different direction from the one originally assumed. Thus, whenever data and biological samples are introduced into a biobank for the first time, the question arises of how precise the consent of the data subject needs to be.

[Rz 5] A number of different solutions have been proposed, and some are already being applied. Depending on the focus, these solutions give priority to either the autonomy of the donor (in the sense of the right to informational self-determination) or the research interests for practicality and implementability of the research projects. The UK Biobank has adopted the principle of a blanket consent, which means that the data subject provides a single authorisation for the data to be used for all future research activities. The German National Ethics Council has also indicated its support for a blanket consent system for biobanks, and does not have any ethical reservations provided that that consent can be withdrawn at any time<sup>7</sup>.

[Rz 6] In its draft guidelines on data biobanks, the Swiss Academy of Medical Sciences (SAMS) also comes out in favour of a blanket consent as long as the standards required by the guidelines are met<sup>8</sup>. A blanket consent would make it possible for the biobank to use samples and data for every type of research – even for research which was not foreseeable at the time consent was given.

[Rz 7] This argument is flawed because failure to restrict the scope or the duration of the consent runs counter not only to the data protection principle of proportionality, but also to the principle of transparency and purpose limitation. It is also questionable whether it is at all possible to obtain informed consent from an individual who consents to something he/she is not aware of yet. What happens if, within the context of a new research project, data are passed on to third parties? Furthermore, civil law protects personality rights by prohibiting any commitment that is made for an indefinite time or purpose<sup>9</sup>. Providing a blanket consent for an unspecified research project and an indefinite period is in my view – even if we assume that it is done for altruistic purposes – simply out of the question. Others argue that that consent should be explicit, that it be tied to a particular research project, and that it be obtained anew if the scope of the original project is either exceeded or modified.

[Rz 8] It is not surprising that researchers feel that this form of consent is excessively bureaucratic, unsatisfactory and could jeopardise research projects. I believe that the only way to strike a balance between these interests is to find a solution that lies somewhere between these two positions. One way of obtaining informed consent is to allow donors to choose whether they want the data to be restricted to a particular project or used for a clearly defined and specific area of research. Donors should also have the possibility of excluding their samples from certain types of research. In either case, it is essential that the donor be able to withdraw consent at any time. This would protect the informational self-determination rights of the data subject.

[Rz 9] In the United States they have developed a new approach. A firm called First Genetic Trust combines informed consent with the latest information technology and provides research institutes with an information platform for the storage of donor data<sup>10</sup>. So-called online dynamic consent is provided over the Internet. Donors are sent information via the web about the state of research and the whereabouts of their data. However, I have my doubts as to whether donors can be expected to understand the information that is sent in this manner without a doctor acting as an intermediary.

### **III. Anonymization and encoding of donor data and samples**

[Rz 10] Once donor data and samples have been entered into the biobank, the next question that needs to be addressed is the anonymization and encoding of the data. If the data and samples are anonymized, it is impossible to attribute them to an individual donor. Even if anonymization of samples is the preferred solution from a data protection perspective, there are certain areas of research where it is important for the donor to be given the results (e.g. in pharmacogenomics). Just as data subjects have the right not to be informed of the results of the research (the so-called right not to know), they also have the right to obtain the results that concern them directly (so-called genetic feedback). Moreover, donors must be able to withdraw their consent to participate in a research project at any time or to invoke their right to be informed. If data and samples are totally anonymized, these rights can no longer be upheld. In other areas of research (e.g. in fundamental research or in simple genome sequencing), it is not necessary

to inform donors of the results, as the research does not aim to reveal any individual information. In such cases, both samples and data should be anonymized from the very outset.

[Rz 11] In cases where anonymization is not appropriate due to donor rights and research goals, the pseudonymization of data, preferably double coded, must be assured in all circumstances. The point here is that coded – and even double coded – data still allow meaningful research to be carried out. Encoding should be done right at the beginning. The encryption key should be deposited with an independent external body (so-called custodians of the data banks), particularly in the case of highly sensitive data. The re-identification of the donor should only be authorized in exceptional and well-defined cases. There can be no justification from a data protection perspective for samples and data for research purposes to be held with the name of the donor without anonymization or pseudonymization.

[Rz 12] The anonymization or encoding of the samples is of great importance, particularly in view of the growing internationalization of research. For example, the UK Biobank is to grant access to both private and public research institutes worldwide. Only a sufficiently sophisticated encoding system can ensure that the desired level of data security will be maintained. The SAMS guidelines mentioned above say that the encoding of samples and data should be standard procedure. Only in exceptional and well-founded cases may data and samples be used in non-encoded or non-anonymized form. The draft guidelines as well call for the encryption key to be held by an external custodian. The guidelines are silent on whether the encoding should be done in simple or double form. In my view, encoding should be adapted to the sensitivity of the data processing and the danger of an infringement of the donor's personality rights – the more delicate the data being researched, the higher the requirements in terms of data encoding.

#### **IV. Data access by third parties**

[Rz 13] A further problem from a data protection perspective concerns the transmission of data or samples to third parties outside the research environment, e.g. the authorities, insurance companies or employers. A broad-based biobank containing personal data and tissue samples allows the identification of individuals based on their DNA profiles and is thus of interest not only to the police, but also, in view of possible pathological examination results, to insurance companies and employers. Access to such data by third parties is not in the realm of science fiction, as two examples from Sweden have proven. As part of its investigation of the murder of Sweden's Foreign Minister, Anna Lindh, the public prosecutor's office asked for particular tissue samples from the national neonatal database, which has been run for the last 30 years, to be made available in order for a DNA comparison to be made. The samples were indeed handed over to the public prosecutor's office, in clear violation of the Swedish Biobank Act, which states that all blood samples collected may only be used with the consent of the person concerned or in pseudonymised form for research purposes. When in January of this year requests were received by the national neonatal database to permit the identification of tsunami victims, Swedish Parliament adopted a special law on the matching of DNA profiles which has been limited to an 18 month period.

[Rz 14] However useful access to the national neonatal database may have been in both cases, they demonstrate just how easy it is to obtain a special authorization or the adoption of legislation in order to gain access to, and thus misuse, a biobank. The same problem arises with insurance companies and employers. If research is being done with the help of a biobank which contains the results of health tests, there may be an interest in gaining access to such data before insurance cover is provided or a job offered.

[Rz 15] Access by third parties needs to be considered from two different angles. The first relates to direct access by the authorities, insurance companies and employers. Data processing in biobanks must be subject to the purpose limitation principle otherwise it is not covered by the donor's consent. Thus, third parties who are not involved in the research project should basically have no right of access to biobanks. However, as this ban can be circumvented by a decree or a law, it is important that the legislator inform the donor if there is a possibility that the data may be used for purposes other than research. The operators of biobanks should make it clear under what circumstances they may be willing or obliged to disclose donors' confidential data. This information should form an integral part of the explanations provided to donors before their consent is sought.

[Rz 16] Another question that needs to be clarified is the issue of data being transmitted within the context of research projects where donors have claimed their right to be informed of the results of examinations. Under the laws of certain countries, donors might be legally required to disclose the results obtained from their participation in a research project to their insurance company or employer. Here, too, donors must be informed if there is a duty of disclosure.

[Rz 17] In Switzerland's case, for example, legislation due to come into force will prevent employers from using the results of genetic tests. Persons who have a life or disability insurance above a certain sum, as well as persons who take out complementary health insurance, can be required under certain circumstances to inform their insurance company of test results before cover is granted<sup>11</sup>. It is important that donors' attention is drawn to any national regulations which could require them to disclose personal data.

## V. Building trust through a research confidentiality policy

[Rz 18] «Trust is a bank's most important capital». This advertising slogan applies equally to biobanks. In many countries confidential data is protected by special professional secrecy requirements. However, these usually fall short of the mark, as they primarily concern specific professional categories such as doctors. Molecular biologists and biochemists also work with biobanks, and they, like the biobank operators, should be subject to an obligation of professional secrecy. A possible solution might be to create, in addition to the doctor-patient privilege, a statutory professional secrecy requirement for researchers through a research confidentiality system. Research confidentiality would be designed to serve the interests of donors by protecting the confidentiality of data and samples from certain persons, institutions or the general public. Research confidentiality would make it much more difficult for third parties to gain access to data.

[Rz 19] Increasing donor trust in biobanks could prove especially important in cases where biobanks are established based on a broad consensus of the donors. Where donors' consent is granted not for each specific research project, but for an entire area of research, it is important that special safeguards be put in place in order to protect their personality rights. Enforcing research confidentiality could strengthen trust in biobanks and personality protection where a broad consensus is given. The introduction of a research confidentiality policy which bans the use of data and samples that are not research-related would not mean, in my view, that the explicit consent of the donor becomes obsolete. Enshrining the principle of research confidentiality in a law cannot replace the consent requirement.

[Rz 20] Research confidentiality only makes sense if the unauthorized disclosure of personal research data is prosecuted. Switzerland already imposes a legal ban through the so-called professional secrecy in medical research<sup>12</sup>.

## VI. Conclusions

[Rz 21] The establishment of biobanks for research purposes raises new challenges worldwide for the legislator in terms of protection of donors' data and personality rights. When weighing up the different interests in connection with biobanks, the interests of research cannot be dismissed; but by the same token, the risks of infringing donors' personality rights must not be trivialised either. When assessing the issues of consent, data use and access rights by third parties, it is important that the sensitivity of collected donor data be included in the equation. Research interests must be taken seriously, but not at the cost of donors' rights, and in particular without violating data protection principles. Although at first sight it might seem that there is very little invasion of donors' privacy when biobanks are set up, the picture regarding the risk to their personality rights is far from clear-cut when considered from today's perspective.

[Rz 22] For legislators in most countries today biobanks still remain *terra incognita*. Switzerland's SAMS has taken a first step towards setting limits for the use of biobank data. However, the discussion is far from over – on the contrary, it has only just begun.

Dr. iur. Claudia Mund holds the position of a Legal Adviser at the Swiss Federal Data Protection Commissioner (SDPC), where she is primarily concerned with human genetics and related data protection issues.

German version of this article:

Biobanken – Datenquellen ohne Grenzen?

---

<sup>1</sup> [www.decode.com](http://www.decode.com); see also the article of Sigrún Jóhannesdóttir on the Icelandic database.

<sup>2</sup> [www.geenivaramu.ee](http://www.geenivaramu.ee)

<sup>3</sup> [www.ukbiobank.ac.uk](http://www.ukbiobank.ac.uk)

<sup>4</sup> [www.popgen.de](http://www.popgen.de)

<sup>5</sup> Swiss Federal Act on Data Protection of 19 June 1992 (DSG, SR 235.1); Council of Europe Convention on the Protection of Individuals with regard to Automatic Processing of Personnel Data (ETS No. 108); EU Directive 95/46/EU on the Protection of individuals with regard to the free processing of personal data and on the free movement of such data (AB No. L 281, 31 ff.).

<sup>6</sup> The issue of how to proceed with existing biobanks is not considered. To do so would have exceeded the scope of this paper.

<sup>7</sup> Cf. German National Ethics Council, Biobanks for Research, opinion of 17 March 2004, Berlin 2004, p. 55 ff., especially p. 61 f. ([www.ethikrat.org/\\_english/publications/Opinion\\_Biobanks-for-research.pdf](http://www.ethikrat.org/_english/publications/Opinion_Biobanks-for-research.pdf)).

<sup>8</sup> The guidelines can be found at [www.samw.ch/docs/Richtlinien/d\\_RLBiobanken.pdf](http://www.samw.ch/docs/Richtlinien/d_RLBiobanken.pdf) (in German).

<sup>9</sup> Cf. for Switzerland Art. 27 ff. of the Swiss Civil Code of 10 December 1907 (ZGB, SR 210).

<sup>10</sup> [www.firstgenetic.net](http://www.firstgenetic.net)

<sup>11</sup> Cf. Art. 21 ff. and Art. 26 ff. of the Swiss Federal Law on the Genetic Testing of Humans of 8 October 2004 (GUMG; BBl 2004, 5483 ff.; [www.admin.ch/ch/d/ff/2004/5483.pdf](http://www.admin.ch/ch/d/ff/2004/5483.pdf)). The law is due to come into force by mid 2006.

<sup>12</sup> Art. 321bis of the Swiss Criminal Code of 21 December 1937 (StGB; SR 311.0)

Rechtsgebiet: Datenschutz

Erschienen in: Jusletter 3. Oktober 2005

Zitiervorschlag: Claudia Mund, Biobanks – Data Sources without Limits?, in: Jusletter 3. Oktober 2005

Internetadresse: <http://www.weblaw.ch/jusletter/Artikel.asp?ArticleNr=4211>