Public Health Genomics

Public Health Genomics 2012;15:293–302
DOI: 10.1159/000336541

Published online: June 20, 2012

From Protection of Privacy to Control of Data Streams: A Focus Group Study on Biobanks in the Information Society

K. Snell[a] J. Starkbaum[b] G. Laußc A. Vermeer[c] I. Helén[a]

[a]Department of Social Research, University of Helsinki, Helsinki, Finland; [b]Life Science Governance Research Platform, University of Vienna, Vienna, Austria; [c]Faculty of Theology, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany

Key Words
Biobanks · Control · Data streams · Focus groups · Information society · Privacy

Abstract
Most people in Europe do not know what biobanks are. In this study, public perceptions of biobanks and collection of genetic and health data were analyzed in relation to other technologies and digital networks where personal information is compiled and distributed. In this setting, people contextualized biobanks in line with their daily experiences with other technologies and data streams. The analysis was based on 18 focus group discussions conducted in Austria, Finland and Germany. We examined the ways in which people frame and talk about problems and benefits of information distribution in digital networks and biobanks. People identify many challenges associated with collection of personal data in the information society. The study showed that instead of privacy – which has been the key term of bioethical debates on biobanks – the notions of control and controllability are most essential for people. From the viewpoint of biobanks, issues of controllability pose challenges. In the information society, people have become accustomed to controlling personal data, which is particularly difficult in relation to biobanks. They expressed strong concerns over the controllability of the goals and benefits of biobanks.

Introduction

We analyzed public perceptions of biobanks[1] [1] and biobank networks and how they are contextualized in relation to everyday information sharing practices and transactions in the Web, which we call data streams. We also carefully assessed how focus group participants in Austria, Finland and Germany think about sharing personal information in different modes of digital interaction (e.g. public registers, Internet banking and shopping,

[1] Biobanks are social and technical entities that retrieve, store, exchange tissues samples and associated medical (e.g. diagnosis and health records) and lifestyle information, or register data. They vary in size and organization from large national collections covering the general population to specialized, small-patient-sample collections connected to a hospital clinic or a research laboratory [1].

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1662–4246/12/0155–0293$38.00/0

Accessible online at:
www.karger.com/phg

Karoliina Snell
Department of Social Research, University of Helsinki
PO Box 18
FI–00014 Helsinki (Finland)
Tel. +358 9 1912 3969, E-Mail karoliina.snell@helsinki.fi

social media) as compared to donating a tissue sample and related health and lifestyle information in biobanks. Our data consist of 18 focus group discussions conducted in these 3 countries in 2011.

Today, people know about the increasing collection, storage and distribution of personal information. They have became aware that they often share large amounts of personal information in global information and communication technologies (ICT) networks as individual users of digitalized shopping, banking, information seeking, and other services, where they leave traces of their personal activities and preferences on the Web. Moreover, they know that databases compile massive amounts of data and make those data readily searchable. Hence, the collection of personal information in public and private digital databases has become more intense with the development of technology for data storage, handling and distribution.

We contextualized biobanks and biobank networks as instances in which personal information is aggregated and distributed in digital networks. The view of biobank activities as streams of biological information is congruent with recent developments whereby biobanks are moving from the national to the international scale. Current biobank projects such as the Public-Population-Project in Genomics (P3G) and the European Biobanking and Biomolecular Resources Research Infrastructure focus on the circulation of biological and personal information, since their purpose is to create technical and governance frameworks that enhance and stabilize transactions between depositories of tissue samples and health data and research institutions [2, 3].

Experts in biobank governance consistently frame people's concerns about the massive collection, distribution and utilization of their biological, medical and lifestyle data as privacy concerns [4–24]. In our study, we examined whether or not privacy indeed plays such an important role in reasoning by people who are not familiar with those debates on ethical, legal or sociological levels. Our analysis focuses on the ways in which participants in the focus groups frame and talk about problems of information provision in digital networks and biobanks. We were interested in people's views about what portions of data streams [25] are controlled by whom, by what means and for what purposes. In addition, we looked at how their viewpoints and willingness to share data differ depending on the ICT context. Finally, we took note of how participants express their needs for control of uses of and benefits from the data they share in the digital universe.

Biological Data in the Digital Universe and the Relevance of Controlling Data Streams

In general, the question of data control seems pivotal to how people feel about the circulation of personal data in cyberspace. According to Eurobarometer surveys, people are aware of security problems on the Internet and are cautious about sharing their personal information [26, 27]. However, people's attitudes seem ambivalent. Three out of 4 (74%) respondents agreed that disclosing personal information is an increasing part of modern life [28], and a majority of 58% felt that it is safe for them to carry out transactions over the Internet [27]. At the same time, 9 out of 10 respondents (86%) said that they avoid giving out personal information on the Internet as much as possible [27]. Differences in attitudes between countries were also considerable: while in Finland only 5% of respondents were 'very concerned' regarding organizations that hold personal data, the share went up to 65% in Germany and 70% in Austria, while the EU-27 average was 34% [26].

With health-related data, European attitudes toward privacy and the protection of personal information were ambiguous as well. Health data was framed as 'personal' by 3 out of 4 survey respondents (74%), and the vast majority of all respondents stated that they have not disclosed personal health information in different ICT contexts [28]. Nevertheless, that people see medical data as personal does not mean they would not provide their data to biobanks. A Eurobarometer survey shows that almost half of the respondents (46%) could conceive of providing information to a biobank. In the 3 countries involved in our analysis, the willingness to do so was 67% in Finland, 42% in Germany and 35% in Austria [29]. The results are congruent with empirical studies that show that privacy issues do matter for people, although these issues do not necessarily prevent them from participating in biobank research [17, 30–32].

In order to explain this, we complemented this quantitative data with focus group analysis and looked closely at the ambivalences and variation in people's views on privacy and data control. These issues are especially interesting in the domain of health-related data, genomics and biobanking, since these areas seem to have a special relationship to privacy debates, as we shall see.

The idea that individuals must be protected against involuntary disclosure of personal genetic information was introduced in the wake of the Human Genome Project [4]. In the context of the Icelandic Health Sector Database Project, this notion of 'genetic privacy' became a

topic that created a long-lasting relationship between biobank projects and bioethical debates. In 2004, the Icelandic Supreme Court ruled that offspring have the right to refuse the transfer of medical and genetic data of deceased parents to the Health Sector Database, on the basis that these records contain information about the offspring [33]. The ruling gave impetus to an intense academic and administrative reassessment of bioethical principles such as informed consent and privacy and their impact on the governance of human genetic databases [34]. Privacy became a dominant bioethical topic in the literature on biobank governance, the gist being that information providers had to be protected from illegitimate access to and deployment of personal data, which were deemed intrusions into their personal sphere [14]. However, in bioethical discussions, privacy has been predominantly considered a moral principle to be balanced against the prospects of future improvements in science and health care for the good of individuals and the population at large. The assumption is that the future health of individuals and populations depends upon research endeavors that need information supplied by patients and ordinary citizens, in order to develop new preemptive measures, diagnostics and treatments. Balanced against future health, theoretical privacy concerns are often eclipsed by the pragmatic needs of medical research and have only very limited practical implications despite such debates.

In discussion on bioethics and governance of biobanks, privacy and its bioethical twin, informed consent are considered problematic or even futile by scholars [16, 35, 36]. This view is congruent with the notion that the development of decentralized ICT networks for global circulation of information has transformed the issue of privacy to be more complicated and even harder to define. Scholars in law, moral philosophy and social sciences have long struggled to define privacy clearly and accurately for use in the increasing dissemination of personal information in cyberspace. If any agreement has been reached on this point, it is that 'privacy' is a multidimensional and equivocal concept in law, moral philosophy, and social sciences, and that its meaning and value are dependent on the context in which it is applied and used [37]. In fact, experts seem to argue that the term 'privacy' cannot be understood as an unequivocal concept, but rather is 'a set of family resemblances' [37]. The division between private and public spheres has always been profoundly blurred [38], and it seems that living in a networked society [39] means that the distinction becomes even less important than it used to be. Despite this ambiguity, many scholars embrace privacy as essentially valuable to personal integrity and as a focal point of personal rights [40].

Given this profound conceptual ambivalence of privacy in the digital age, we tried to understand this problem in its technological and social context [37, 41]. This was the task we pursue in the following analysis.

Methodology and Focus Group Design

The starting point for our research was to contextualize biobanks to different data streams and conduct focus groups with this theme. Focus groups are organized group discussions of 5–10 people. A moderator leads the group with the intention to keep the discussion on track [42, 43]. We sampled our groups specifically to create some degree of internal homogeneity to facilitate group discussions that reflect on given societal structures [44]. In total, we organized 18 groups in Austria, Finland and Germany. This sample provided a diverse scope of countries, showing significantly different attitudes toward biotechnology [29]. We conducted 6 groups in each country with lay people between the ages of 19 and 75, and with a total number of 119 participants. In this article, we did not contrast people's views between these national research sites, but observed general patterns of meaning between different technological contexts and data streams.

We developed a common design that was applied in each country. To improve comparability, we created a script that was used for all groups. This script was designed openly to provide space for the groups to develop their own structures of meaning. We started out by asking: 'When you think about your own life, there are different types of information and data that you give away or that are gathered for different purposes. What information and purposes can you think of?'

Our approach was not to talk about privacy as an issue, but to ask people about data streams in general, let them choose examples and allow them to narrate their individual experiences. As previous research has shown, many people in Europe are not aware of biobanks at all [29]. Therefore, we were prepared to introduce this topic to the participants during the focus group discussions. For comparability, we had decided to introduce some types of data streams to all groups, if they had not already been discussed: social media, electronic health records and biobanks. Most often people raised and discussed the first 2 examples intensively on their own. Biobanks was often the only topic that was introduced by the moderator.

Focus groups, in a sense, offer an artificially created possibility to observe the dynamics and structures of decision-making under laboratory conditions, although there is always some influence due to the different local moderators. This together with group effects, such as social desirability, apply to everyday interactions and have been considered in the analysis [45–47]. The data we present and analyzed worked toward an understanding of how people orient themselves towards more or less unfamiliar terrain. This should help develop ideas about how potential biobank participants reason together and develop ideas and storylines about biobanking, for which they did not possess firsthand experience

or expert knowledge. We also discovered how they rely on previous personal experiences and how they differ, using a comparative qualitative analysis approach.

All focus group discussions were recorded with the consent of the participants and transcribed in each original language. Due to the open concept of data collection, we performed inductive open coding [48, 49] with the first groups and then developed a common coding structure that was applied deductively. We performed a computer-assisted qualitative data analysis that facilitated data management and international exchange and control [50]. The following sections present our findings and analysis with exemplary quotations. Quotations are preceded by the country code (AUT, FIN or GER), focus group number (G1–G6) and the participant number in the group (1–8).

The Challenges of Data Streams

Focus group participants easily identified a number of situations, technologies and applications that require the collection of information from people, information that is then shared in digital form, for example, social media, consumer cards, online shopping and banking, health records, and public registers. In talking about these data streams, people recognized and contemplated different challenges related to the gathering, storage and distribution of information. Many of these challenges were connected to the flow of personal information, but interestingly, privacy was seldom explicitly articulated in people's narratives. In fact, the word 'privacy' was mentioned only a few times in each group, irrespective of country. Thus, privacy did not seem to be the most relevant question related to data streams for people, yet it was not unimportant. Sharing and protecting private information was not usually discussed in terms of intrusion of privacy. Instead, the value of privacy and personal information depended on the technological and social setting, which means that the distribution of information is often actively regulated and deployed in various contexts [51]. Thus, the topic veered away from personal privacy to issues of controllability of data streams.

From our focus group discussions, we have discovered 2 major challenges to controlling data streams. First, people recognize that an increased amount of data is nowadays being collected at various occasions, and this leads also to increased possibilities of linking personal data and information. As a consequence of this complex situation, people identified difficulties in controlling data flow and referred to the 'uncontrollability' of information. People were interested in and puzzled about the ways data streams can or cannot be controlled by individual persons and institutions. They contemplated personal strategies of sharing and protecting information and questioned what kind of data is being gathered, who has access to it and how it is regulated. The second challenge of controllability concerns the deployment and benefits of the data streams: people wondered whether the benefits and purposes of data streams are controllable and how the data streams influence future developments.

The fact of increased collection and circulation of personal data were acknowledged right from the beginning of the focus group discussions. This perception was strongly associated with the rapid developments in the field of ICTs and with the fact that we live in information societies. The general information overload and complexity provoked much discussion about the control and uncontrollability of data streams, and people mainly agreed that much personal data are already 'out there'. Living in the information society was also associated with leaving traces. People talked about digital tracks and how our 'digital footprints' are impossible to erase. People recognized that modern technology prevents us from 'hiding from Big Brother' or escaping from the collecting and aggregation of the personal data:

> FIN (G6) 5: It is like that that you cannot do anything without someone knowing it – unless you are in the forest. How they register information: when you get on a bus, and in shops, they can do research by following [electronic] cards.

People concluded that they can do little about this development. To a large degree, they accepted that uncontrollability of information streams and chaos are part of living in the information society. They assumed that much information is being collected and circulated about them, but they admitted that they do not know exactly what data are gathered and stored, and where and how the data are used.

> GER (G3) 4: This is somehow nontransparent. I think the stupid thing about it is that I get the feeling of becoming a bit more transparent myself. But what is stored is not any more transparent to us. It is like in an interrogation room where there is this pane, where only one side can see, and the one who is sitting inside, who is interrogated, cannot figure out exactly what is happening on the other side.

Thus the possibilities of controlling data streams were important topics of discussion – how an individual can control data streams and what kind of institutional control there is for data streams. Many of the identified risks were linked to the society and future developments. It was often not specific data or technology that caused uneasiness in the discussions, but the broader aims and general

developments all together. Participants were aware that they are dealing with a future, which is in large part hard to predict. This concern is particularly present with biobanks. Another heavily discussed topic dealt with the expected benefits and their distribution. People fear that some might gain, while others do not, and that they can do little about it.

AUT (G4) 4: Who acts? Who collects the data? Who does what with it? Those are those who act. We can only move up; they are ahead of us. But why are they ahead of us, and who is it? And do we need all of this?

In the following we show how people deal with data streams and identified challenges in their personal lives.

The Controllability and Uncontrollability of Data Streams

The way people expressed themselves in the course of the focus groups supports the view that the perceived problems with controllability and uncontrollability of data streams need somehow to be dealt with. They simultaneously acknowledged uncontrollability and stated a strong demand for control. Here we present 4 dimensions of controllability that people referred to in the discussions. These different dimensions of controllability coexist, overlap and were used in parallel in the discussions, although it became evident that the 4 dimensions of controllability are manifested to different degrees in relation to different technologies.

First, many people expressed a strong belief in their own ability to control what information they share and what part of the data streams they participate in. Second, people either assumed that many of the data streams are controlled by law and regulations anyway, or they had a desire for these to be institutionally controlled. The third common reaction to the tension between the widespread existence and uncontrollability of data streams was simply to face the loss of control or to stop showing personal interest in controlling the data streams. Here we encountered an attitude of frustration. Fourth, people turned their focus from control of sharing and gathering of their personal data to controllability of the purposes and benefits of the data streams. That is, they substituted their desire to control every bit of personal information by a desire to have a say on the contexts and manners in which the provided information is utilized.

Active Self-Regulation of Data Streams

Almost everyone believed they can control, to some extent, data flow by regulating what data and to whom they give them. To many, the amount of expected personal control – the type of data shared, to whom and with what conditions – creates confidence when they engage with a technology. The discussions about social media and Internet use were in many cases dominated by narratives about different options for controlling data streams and access on the basis of individual strategies and hence individual competence. Other studies show similar results [52, 28]. While people seemed aware that any form of individual control is limited, and that corporations like Facebook and Google have the upper hand, there was still broad confidence about different modes of control in this field. These modes of control include, for example, adjusting technology-based settings that regulate access, as with who can see your photos on social network sites. But the most direct way of using personal control is to regulate the quantity and quality of disclosed data – in certain technological contexts, many have a special type of data they want to protect, such as photos, family connections, phone numbers, and Social Security numbers.

FIN (G3) 1: It is up to you what you give there. You put on the Net what you want others to know.

GER (G4) 1: Sensitive data, we all agree on, are health data, as we discussed and everything regarding our finances. Furthermore, when we have to fear that it could have consequences for our professional life.

Stripping data of personal identifiers to make it more anonymous or using multiple avatars, profiles or even identities was perceived as another reasonable strategy to protect personal data.

AUT (G6) 4: I have recently talked to a mother of a, don't know, a 12-year-old. She said her daughter and her friends are, even if it is not allowed, all on Facebook – it is like this. She is very happy that they all at least do not use their own names, because, I mean, you indeed write one or another stupid thing on it [Facebook].

Second profiles or virtual personalities are ways to deal with data chaos. Self-regulation is not only an active means of protecting privacy or regulating what data you share, but also a way to control the person you want to present yourself as.

FIN (G1) 6: It is also about constructing digital footprints, that you grow your own media persona or personality on the Net. In fact, I think that you can fake on the Net. You can build yourself a persona, for example, through blogs.

In other cases, a person's ability to control data streams was connected to strict conceptions about data protection and a refusal to participate. People can mistrust the insti-

tutions that handle data streams, or they can question in general the relevance of data sharing and collecting.

GER (G1) 5: I give out my personal data only if I must. Therefore, I do not participate in anything.

In talking about biobanks, people expressed their options of personally controlling data streams very differently from, for example, social media. On the one hand, people anticipated that the amount and type of data given to a biobank cannot be personally regulated. They also recognized that any form of continuous individual interference that regulates data access is not perceived as an option. On the other hand, people wanted to act responsibly – to have control over the use of their personal data – and considered during the discussion whether this would be possible with biobanks.

AUT (G5) 2: There must be some criteria like informed consent, where people have to agree on, knowing that data are being collected. Such basic rules or that you can always say: no, I don't want to give my tumor for research purposes.

Because the modes of control that seem to work with technologies such as social networks are not applicable for biobanks – apart from a refusal to participate – people regarded their demand for anonymity as viable. However, total anonymity in practice is generally not feasible for biobanks, as they rely on personalized and longitudinal data sets [53].

Institutional Regulation and Control

Most people agreed that any form of individual controllability is limited and that some external regulations are expected for all technological contexts. And as described, controlling access and structure of data streams was often associated with personal competence, although it was a recurring topic that not everyone has sufficient competence for self-regulation. Since there is no universal solution on the individual level, other modes of control were demanded, especially for those who are perceived as needy or helpless – such as children or the elderly. People also argued that there should be more education in how to handle these kinds of data flows in order to gain better knowledge of new phenomena and technologies. Institutional and external regulation is expected and demanded, but was not extensively debated.

When it comes to biobanks and medical research, people conventionally assumed that appropriate regulations are already in place. At the same time, they pondered the practical relevance or proper functioning of control mechanisms. Nevertheless, people expressed a strong demand for different common control mechanisms.

GER (G4) 6: It has to be guaranteed under any circumstances that there is absolute anonymity and also corresponding licenses.

AUT (G5) 1: As I mentioned, I am not very well informed, but I am pretty sure that there are pretty clear legal guidelines on how data can be used, even if they are anonymous. But there is the possibility of misconduct and breaches of the law. And then there is a scandal. But most of the time this leads to a process of muddling through, and after that everything seems somehow all right again … The problem is that there is a huge gray zone.

So, even if people assumed that appropriate regulation is in place or should be made as good as possible, they sensed that data protection issues operate in gray zones in which no legal certainty exists. Participants often mentioned that no control mechanism or data security system can ever be 100% safe. Some emphasized monitoring the practical ways in which people handle data streams, such as researchers and medical staff who deal with the information. So, although we observed quite high levels of trust in the existence of regulation and its general quality, there was doubt about the reliability of the human factor in systems in which practices of data use are nontransparent and in which data users can stay anonymous.

FIN (G5) 2: I was applying for income support, and I have been working in a place where you can access this database. I could have applied, but I realized that my former coworkers – there are couple of nosy ones – go to the database and look to see if FIN (G5) 2 has applied for income support. So I didn't apply.

The burden of control and regulation with biobanks and health data was mostly placed on official public institutions or the government. In general, publicly governed data streams were considered to be more trustworthy than private ones. Trust is mentioned as a basic resource for any kind of data exchange. In a way, trust is the social medium in which data streams can grow, multiply and replicate themselves.

AUT (G1) 1: Data provision has very much to do with trust. In principle, I provide data to someone I trust.

Commercial impacts were identified as potential threats to controllability. This was particularly apparent in relation to health data and biobanks, where commercialization and involvement of the pharmaceutical industry created worries about research aims and distribution of benefits.

FIN (G1) 1: I would draw the line there that I would not give mine [data] to the pharmaceutical industry. I would definitely want it to go to scientific research.

Commercial data streams seemed to be much harder to control than streams more directly associated with public

authorities. Similarly, commercial players were regarded as harder to control, and public institutions were considered to be the proper domain for research. At the same time, this dichotomization affects the levels of protection that people demand from different institutions: they do not expect the same amount of transparency and controllability from data streams that are perceived as commercial. This was the case with pharmaceutical companies, bonus-card providers and Internet service providers. The public biobanking infrastructure lies on the other end of the spectrum and was therefore confronted with higher demands for responsible behavior and institutional control.

Accepting Loss of Control

People thought that self-regulation and institutional control of data streams have their limits and that through these means only some of the challenges of control can be dealt with in a reasonable manner. Therefore, many surrendered to uncontrollability and acknowledged that there is little one can do about the data chaos except to accept the situation as it is.

FIN (G1) 3: But surfing on the Net, it means that you always leave traces and that is impossible to control. In fact, I have a strong view that all in all, controlling information today is more or less impossible.

Participants rationalized that most of the data gathered are more or less irrelevant, useless, already otherwise easily available, or harmless. If the data as such are not perceived as having the potential for misuse, then they don't need to be controlled. The same logic applied to individual persons. If the individual is not considered an important or famous figure, the possibilities for abusing their data are there, but participants did not believe that anyone would benefit from misusing the data. People considered themselves to be ordinary and to be living a life that can withstand scrutiny and publicity. In some cases, people framed those that act deviant as being guilty and having something to hide.

AUT (G2) 1: I have to live a correct, serious and proper life. So I don't have to be anxious about the ways my data are used.

FIN (G2) 5: The trust is somehow increased by the fact that why someone would be interested in my life so much that they would want something. I would understand if I were a boss of a big computer firm and would have big secrets there. But an ordinary person …

Many people expressed that they do not have the time, energy or desire to perform control of their data streams. Understanding the settings of Internet services, reading privacy policies and searching for advice on how to deal with data streams were seen as time-consuming, and people talked about the strain of accomplishing these tasks. This attitude can also be interpreted as an example of deliberate ignorance [54]; people hoped and expected that, for example, data protection is taken care of by experts. With biobanks, it has been demonstrated that people rarely read informed consent forms [55], which has been explained with the trust that people have toward the medical staff or institutions running the biobank. However, participants recognized that not reading consent forms or terms of privacy is a risk:

GER (G4) 1: The problem is, I think that you inevitably deal much too carelessly with it [data control]. I find myself doing it again and again. If I only think about 2 pages of general terms and conditions or something else – ok, I don't have time now, so we skip it this time, and I don't know, what is going to happen with it [personal information]. This is always a critical point.

This kind of attitude in many cases had to do with the benefits people expect from their participation in data streams. Different forms of incentives influence how the engagement with a certain technology is framed and valued.

Controlling Goals, Benefits and Risks of Data Streams

The expected benefits significantly influence how people perceived data streams with certain technologies and related challenges. Benefits were a strong incentive for them to engage with a technology. Expectations about both individual *and* common benefits influenced the perception of related matters, such as impacts or risks. In the case of social media, consumer cards or many Internet services, for example, benefits were mainly expected at a personal level. The benefits were regarded to be partly controllable or they are at least anticipated.

As with other technologies, these benefits are not an independent issue that people acquire, but they are integrated into people's personal lives. This means that data streams often offer benefits that people can only barely resist. In order not to be excluded, people, for example, engage with social networks despite their awareness of the uncertainties [56]. Many data streams directly serve personal and social needs.

AUT (G1) 1: I think with Facebook it is, at least for me, simply that I know that it bears risks, but my social environment is, on the other hand, so active there that simply my – this sounds totally stupid – but I would miss something from my social environment, simply, because I get invited to events.

But, as stated, benefits are not only personal. In many cases, people identified possible common benefits from data streams, and personal and common benefits were

often intertwined. This was apparent with health-related data streams that include biobanks and electronic patient records. For example, having a centralized national medical record database was seen to benefit the community through efficiency and the individual by potentially saving one's life.

FIN (G4) 4: For example, if you have some medication and you go to a strange locality and get into an accident. It would be good that this patient information would be available everywhere … It would be good if they could see immediately the blood type, for example.

GER (G5) 1: I am open to the collection of data because we need this to make sense of our research. Research is not done for its own sake – research supports also politics and political decision-making. If I can recognize this, then I even don't need to have a personal benefit; then I am completely open to these kinds of things [biobanks].

With health data in general and biobanks in particular, people more likely used narratives that refer to communal benefits. For many people, medical research and developments were strongly related to an idea of improving people's life and society at large, and in this way of perceiving, concerns and risks seemed secondary.

AUT (G2) 1: I am totally for it that these things [biobanks] exist. Because research aims in this direction, and if one single human life can be preserved, extended, more liveable – we all would not sit here with the diseases we have if this was nonexistent.

However, common benefits to be gained in the future were not seen without complications. Focusing on future consequences raised the issue of control to another level. In this respect, people were more worried about whether the data in general will be used in a beneficial way in the future than about the direct misuse of their personal data. People expressed their limited opportunities to control and have an impact on biobanks' aims, and they articulated concerns regarding the possibilities for institutional control over the functioning of research infrastructures.

GER (G3) 4: I find it important especially concerning those biobanks and in general data banks that they have a time limit for using the data. Today we abandon our data, but in 25 years, the framework will change and then they could use these data again.

The risks of biobanks were associated with the uncontrollability and unpredictability of their long-range goals and impacts. In other words, people were worried about the research goals in the long term. They recognized that their data can be stored for decades and that no one can predict the political and social development of the society – for example, the failing of democracy was seen as a threat for controllability. Another specific threat for common benefits of biobanks talked about in all of the 3 countries was commercialization. Biobanks' financial gains are contrasted to common benefits. Commercialization and globalization make potential benefits remote and less tangible for people.

Conclusions

Concerns about the collection, storage, distribution, and utilization of biological data have been framed in terms of privacy concerns in expert debates on biobank governance. In our study, we examined whether privacy is such a fruitful concept for understanding people's attitudes and concerns about biobanks. People's opinions on biobanks have been studied before from a number of viewpoints – information feedback, informed consent and general acceptance [29, 31, 32, 54, 57, 58] – but our aim was to put people's attitudes in the context of their daily interactions in an information society. We wanted to comprehend how people respond to different data streams and to the challenges they present.

The findings of our focus groups, conducted in Austria, Finland and Germany, provide insight into the way people conceptualize control of personal data streams in general and in regard to biobanks in particular. We have concentrated here on similarities across the 3 countries – challenges and issues that are present in all of them despite differences in attitudes that have been found in other research. The most compelling findings center on the issue of controllability and related challenges. People were worried about an increasing amount of personal data being collected in a number of nontransparent instances. At the same time, people were concerned that their data are not used in accordance with their preferences and that the benefits are not distributed in a justifiable manner. We identified different ways of approaching these challenges of controllability: active self-regulation of data streams, demanding institutional regulation and control, accepting the loss of control, and finally, controlling goals, benefits and risks of data streams. People's concerns with biobanks did not reflect the privacy debate. In fact, privacy seemed not to be the major anxiety associated with biobanks for 2 reasons. First, people expected that biobanks have taken care of data protection as far as possible. They reasoned that biobanks do not want to risk mishandling people's private information, but they also believed, justifiably,

that no system is absolutely secure. Second, people focused more on the goals of biobank research, and their concerns were related to the long-term development of research, society and the political system. These future developments were seen as potential threats that are not controllable by individuals or even by authorities [59]. People expressed a desire to have control over their data and simultaneously they acknowledged the uncontrollability of data streams in the information society. Different types of data streams and technologies have their own particular characteristics. Controllability of personal data was therefore perceived to be different in the case of social media than with biobanks. People actively use self-regulation in sharing and protecting personal data on the Internet, whereas with biobanks it was understood to be more difficult.

From the viewpoint of biobanks, issues of controllability pose many challenges. With other technologies, people have become accustomed to self-regulating data streams [32]. Focus group participants applied this idea to biobanks as well. Demand for controlling the use of personal data and receiving individual results is on the rise [60]. But the desire for control was not limited to personal results and privacy. People were concerned over the research goals and aims of biobanks and how they will be governed in the future. Therefore, public discussion about biobanks needs to extend the issue of privacy to wider debates about the future of research and its societal consequences. The most gripping challenge for international biobank research infrastructures comes, however, from the perceived threats of globalization and commercialization. Local and national public biobanks were regarded to be relatively trustworthy and able to have some control over the data streams. With international, large-scale public-private research collaboration, participants felt that benefits, risks and data streams become uncontrollable.

Acknowledgements

The research for this article was conducted in the context of the project 'PRIVATE Gen' (Privacy Regimes Investigated: Variations, Adaptations and Transformations in an Era of (Post-)Genomics) which is supported by the German Federal Ministry of Education and Research (BMBF), DLR, Academy of Finland and FFG. The authors would like to thank the research partners and funding agencies that made this research possible. We also thank all the participants who took part in the focus groups in Austria, Finland and Germany.

References

1 Hirtzlin I, Dubreuil C, Préaubert N, Duchier J, Jansen B, Simon J, Lobato De Faria P, Perez-Lezaun A, Visser B, Williams GD, Cambon-Thomsen A; Eurogenbank Consortium: An empirical survey on biobanking of human genetic material and data in six EU countries. Eur J Hum Genet 2003;6:475–488.
2 Yuille M: Infrastructure vital to genome success. Nature 2011;471:166.
3 Knoppers BM, Newton J: Creation of population biobanks: design and conduct. P3G Observatory. http://www.p3gobservatory.org/repositoryRefQuestionnaires.htm.
4 Roche PA, Annas GJ: Protecting genetic privacy. Nat Rev Genet 2001;2:392–396.
5 Laurie G: Genetic Privacy: A Challenge to Medico-Legal Norms. Cambridge, Cambridge University Press, 2002.
6 Wylie JE, Mineau GP: Biomedical databases: protecting privacy and promoting research. Trends Biotechnol 2002;21:113–116.
7 Malin BA: An evaluation of the current state of genomic data privacy protection technology and a roadmap for the future. J Am Med Inform Assoc 2005;12:28–34.
8 Malin BA: A computational model to protect patient data from location-based re-identification. Artif Intell Med 2007;40:223–239.
9 Roche PA, Annas GJ: DNA testing, banking, and genetic privacy. N Engl J Med 2006;355:545–546.
10 Kapp MB: Ethical and legal issues in research involving human subjects: do you want a piece of me? J Clin Pathol 2006;59:335–339.
11 van Veen BE: Human tissue bank regulations. Nat Biotechnol 2006;24:496–497.
12 van Veen EB: Obstacles to European research projects with data and tissue: solutions and further challenges. Eur J Cancer 2008;44:1438–1450.
13 Nordal S: Privacy; in Häyry M, Chadwick R, Arnason V, Arnason G (eds): The Ethics and Governance of Human Genetic Databases – European Perspectives. New York, Cambridge University Press, 2007, pp 181–190.
14 Räikkä J: Autonomy and genetic privacy; in Launis V, Räikkä J (eds): Genetic Democracy: Philosophical Perspectives. UK, Springer, 2007, pp 43–51.
15 Ursin L: Biobank research and the right to privacy. Theor Med Bioeth 2008;29:267–285.
16 Lunshof JE, Chadwick R, Vorhaus DB, Church GM: From genetic privacy to open consent. Nat Rev Genet 2008;9:406–411.
17 Kaufman DJ, Murphy-Bollinger J, Scott J, Hudson KL: Public opinion about the importance of privacy in biobank research. Am J Hum Genet 2009;85:643–654.
18 P3G Consortium, Church G, Heeney C, Hawkins N, de Vries J, Boddington P, Kaye J, Bobrow M, Weir B: Public access to genome-wide data: five views on balancing research with privacy and protection. PLoS Genetics 2009;5:e1000665.
19 Austin L, Lemmens T: Privacy, consent and governance; in Dierickx K, Borry P (eds): New Challenges for Biobanks: Ethics, Law and Governance. Oxford, Intersentia, 2009.
20 Knoppers BM, Abdul-Rahman MnH: Biobanks in the literature; in Elger B, Biller-Andorno N, Mauron A, Capron AM (eds): Ethical Issues in Governing Biobanks. Aldershot, Ashgate, 2008, pp 13–22.
21 Knoppers BM, Abdul-Rahman MnH: Health privacy in genetic research. Politics Life Sci 2009;28:99–101.
22 Knoppers BM: Consent to 'personal' genomics and privacy. Direct-to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent. EMBO Rep 2010;11:416–419.

23 Townend D: Privacy, health insurance, and medical research: tensions raised by European data protection law. New Genet Soc 2010;29:477–493.
24 Bialobrzeski A, Ried J, Dabrock P: Privacy revisited? Old ideals, new realities, and their impact on biobank regimes. Poiesis and Praxis 2011;8:9–24.
25 Hilgartner S, Brandt-Rauf SI: Data access, ownership, and control: toward empirical studies of access practices. Sci Commun 1994;15:355–372.
26 European Commission: Data Protection in the European Union. Flash Eurobarometer 225. Brussels, European Commission, 2008.
27 European Commission: Confidence in the Information Society. Analytical Report. Flash Eurobarometer 250. Brussels, European Commission, 2009.
28 European Commission: Attitudes on Data Protection and Electronic Identity in the European Union. Special Eurobarometer 359. Brussels, European Commission, 2010.
29 European Commission: Europeans and biotechnology in 2010. Winds of change? Eurobarometer 24537. Luxembourg, European Commission, 2010.
30 Allen A: Genetic privacy: emerging concepts and values; in Rothstein M (ed): Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era. New Haven, Yale University Press, 1997, pp 31–59.
31 Hoyer K: Donors perceptions of consent to and feedback from biobank research: time to acknowledge diversity? Public Health Genomics 2010;13:345–352.
32 Tupasela A, Sihvo S, Snell K, Jallinoja P, Aro AR, Hemminki E: 2010: Attitudes towards the biomedical use of tissue sample collections, consent and biobanks among Finns. Scan J Public Health 2010;38:46–52.
33 Merz JF, McGee GE, Sankar P: 'Iceland inc.'?: On the ethics of commercial population genomics. Soc Sci Med 2004;58:1201–1209.
34 Chadwick R, Cutter M: The impact of biobanks on ethical frameworks; in Häyry M, Chadwick R, Arnason V, Arnason G (eds): The Ethics and Governance of Human Genetic Databases: European Perspectives. Cambridge, Cambridge University Press, 2007.
35 Lunshof JE, Chadwick R, Church G: Hippocrates revisited? Old ideals and new realities. Genomic Med 2008;2:1–3.
36 Corrigan O: Empty ethics: the problem with informed consent. Soc Health Ill 2003;25: 768–792.
37 Solove D: 'I've got nothing to hide' and other misunderstandings of privacy. San Diego Law Review, 2007, p 745.
38 Geuss R: Public Goods, Private Goods. Princeton, Princeton University Press, 2003.
39 Castells M: The Information Age: Economy, Society and Culture. Oxford, Blackwell Publishers, 2000.
40 Rössler B: Der Wert des Privaten. Frankfurt am Main, Suhrkamp, 2001.
41 Nissenbaum H: Privacy as contextual integrity. Wash Law Rev 2004;79:119–158.
42 Bloor M, Frankland J, Thomas M, Robson K: Focus Groups in Social Research. London, Sage, 2001.
43 Krueger R, Casey M: Focus Groups: A Practical Guide for Applied Research, ed 4, Los Angeles, Sage, 2009.
44 Mangold W: Gegenstand und Methode des Gruppendiskussionsverfahrens. Aus der Arbeit des Instituts für Sozialforschung. Frankfurter Beiträge zur Soziologie. Mannheim, Europäische Verlagsanstalt, 1960, vol 9.
45 Cicourel AV, Herrero EF: Method and Measurement in Sociology. New York, Free Press of Glencoe, 1964.
46 Hollander JA: The social contexts of focus groups. J Contemp Ethnogr 2004;33:602–637.
47 Smithson J: Using and analysing focus groups: limitations and possibilities. Int J Soc Res Methodol 2000;3:103–119.
48 Charmaz K: Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks, Sage, 2006.
49 Strauss AL, Corbin J: Grounded Theory: Grundlagen Qualitativer Sozialforschung. Weinheim, Beltz, 1996.
50 Lewins A, Silver C: Using Software in Qualitative Research: A Step-by-Step Guide. Los Angeles, Sage, 2007.
51 Petronio S: Boundaries of Privacy: Dialectics of Disclosure. Albany, State University of New York Press, 2002.
52 Waters S, Ackerman J: Exploring privacy management on Facebook: motivations and perceived consequences of voluntary disclosure. J Comput-Mediat Comm 2011;17:101–115.
53 Sándor J, Bárd P: Anonymity and privacy in biobanking; in Lenk C, Sándor J, Gordijn B (eds): Biobanks and Tissue Research. New York, Springer, 2011, pp 213–230.
54 Michael M: Ignoring science: discourses of ignorance in public understanding of science; in Irwin A, Wynne B (eds): Misunderstanding Science? Cambridge University Press, 1996, pp 107–125.
55 Skolbekken JA, Ursin LØ, Solberg B, Christensen E, Ytterhus B: Not worth the paper it's written on? Informed consent and biobank research in a Norwegian context. Crit Public Health 2005;15:335–347.
56 Youn S: Teenagers' perceptions of online privacy and coping behaviours: a risk-benefit appraisal approach. J Broadcast Electron 2005;49:86–110.
57 Wolff K, Brun W, Kvale G, Ehrencrona H, Soller M, Nordin K: How to handle genetic information: a comparison of attitudes among patients and the general population. Public Health Genomics 2010;13:396–405.
58 Lemke AA, Wolf WA, Hebert-Beirne J, Smith ME: Public and biobank participant attitudes toward genetic research participation and data sharing. Public Health Genomics 2010;13:368–377.
59 Hoeyer K: Trading in cold blood? Trustworthiness in face of commercialized biobank infrastructures; in Dabrock P, Tauptz J, Reid J (eds): Trust in Biobanking. Dealing with Ethical, Legal and Social Issues in an Emerging Field of Biotechnology. Berlin, Springer, 2012, pp 21–42.
60 Wallace SE, Kent A: Population biobanks and returning individual research results: mission impossible or new directions? Hum Genet 2011;130:393–401.