Personal data in the life sciences

Helping researchers handle data protection and ethical requirements

BioMedBridges Knowledge Exchange Workshop

Authors: Murat Sariyar[1], Irene Schluender[1], Tom Hancocks[2], Stephanie Suhr[2]

1 TMF – Technologie- und Methodenplattform e. V., Berlin, Germany

2 EMBL-European Bioinformatics Institute, Hinxton, U.K.

Executive Summary: Organised by TMF and the BioMedBridges training team, technical staff, legal and ethics experts met on 30 June 2014 in Berlin for a knowledge exchange workshop to discuss how best to support life science researchers who need to use sensitive data and/or biosamples. Presentations outlined current and proposed EU data regulations, and synergies between existing resources that aid researchers in navigating various local, national, European and international legislation were explored. The discussions reflected the complexity of the issue and highlighted (1) that the different approaches assume different user needs and prior knowledge of the subject, affecting how a service is designed and the usefulness of it is assessed, (2) that there is much potential for collaborations, and (3) that enriched annotations of services (e.g. update status, completeness of information, disclaimers etc.) would increase their value and facilitate their quick assessment by (potential) users. Overall, it became obvious that there is still work to do with respect to providing researchers using sensitive data or samples with truly 'useful' tools that do not require pre-existing, in-depth knowledge of Ethical, Legal and Social Implications (ELSI) or plenty of time to delve into the details. Ultimately, separate resources, maintained by experts familiar with the respective fields of research, may be needed while - in the longer-term - harmonization and ease of use for the researcher are very desirable.

Contents

Participants

Bild, Raffael	Technical University Munich
Bovenberg, Jasper	Legal Pathways
Brandizi, Marco	EMBL-European Bioinformatics Institute
Chassang, Gauthier	INSERM
Chatziastros, Astros	TMF e.V.
Drepper, Johannes	TMF e.V.
Gormanns, Philipp	Helmholtz Centre Munich
Hancocks, Tom	EMBL-European Bioinformatics Institute
Karakoyun, Toeresin	Heinrich-Heine University Duesseldorf
Leinonen, Jaakko	CSC – IT Center for Science Ltd.
Sariyar, Murat	TMF e.V.
Suhr, Stephanie	EMBL-European Bioinformatics Institute
Tassé, Anne-Marie	Public Population Project in Genomics and Society
Verburg, Annemarie	VU University Medical Center
Schluender, Irene	TMF e.V.
Siddiqui, Roman	TMF e.V.

1 Introduction

Ethical, Legal and Social Implications (ELSI) form subject-unrelated challenges for the biomedical researcher that are frequently only regarded as obstacles. At best, ELSI is seen as a means for gaining trust and motivating participation in biomedical research. Therefore, it is especially important to support and guide researchers through these issues which they are not trained for. The interactive BioMedBridges workshop "Personal data in the life sciences" was arranged to discuss this topic and had four main aims:

- to gain an overview of challenges on data protection & privacy in biomedical research
- to get in insight into existing resources and services for aiding the researcher
- to define gaps and future challenges
- to elaborate collaborative solutions.

The primary motivation to organize this workshop was to place the "BioMedBridges Legal & Ethical Assessment Tool[1]" into a wider context, discuss whether the tool is a step in the right direction and to investigate how it may be improved by investigating other approaches.

2 Overview of challenges

2.1 EU data protection regulation: implications for scientific research

Irene Schluender presented the main problems when dealing with privacy and data protection in the research context. As a general requirement, the data subjects have to have control over their data, and one important way to enable such a control is the usage of informed consent. However, the details and broadness allowed for informed consent are intricate issues, especially on the

[1] http://hhu3.at.xencon.de/web/guest/bmb-assessment-tool

European level. The need for harmonization is obvious, not only with respect to informed consent. The European Directive 95/46/EC was issued to enhance ELSI interoperability in Europe. Under this directive, anonymous data are not considered to be personal data and therefore are not protected while every other kind of human data – pseudonymized or not – can only be processed if informed consent of the data donor is in place or there is a legal basis. Different national implementations and new problems generated, for example, by the ever-increasing amount of genomic data and long-term storage of biosamples necessitated further improvement in the regulations. Among others, the following questions were posed during the presentation:

- How to avoid fragmentation of national laws?
- What does informed consent mean in practice?
- How broad can informed consent be?
- Can the “relative anonymity”- concept be used to process non-identifiable parts of the data or is that concept too theoretical?
- Can genomic data and biosamples in general be anonymised at all?
- What is the role of ethics committees and data protection agencies, especially in multi-centered research, and on what basis do they decide?
- How should incidental findings be dealt with?

The current draft of the EU General Data Protection Regulation tries to address some of the problems and questions presented. However, many areas are still not covered and additional problems would be generated by the Regulation based on the current draft. For example, the draft insists on the necessity of more or less specific consent while still allowing exemptions at the national level, even though the hurdle for such an exemption is a little bit higher by requiring concrete proof of the significance of the national interest. Areas such as long-term storage of biosamples or processing of genomic data are not covered. Under this draft, biobanks as an important infrastructure for basic biomedical research cannot be sure whether broad consent would provide a valid legal base for data use and re-use. Overall, it seems that the draft would be a step backwards for biomedical research if it was adopted in the current form.

The trade-off ‘social benefits of biomedical research vs data protection’ is ubiquitous, and the presentation made clear that legislation alone cannot find the right decisions. The research community and ELSI experts have to come together to produce viable solutions and inform the legislation process. The concept of anonymity provides an example for this: legislation does not point to concrete technical procedures for anonymizing data sets or to assess the danger of identity disclosure that is possible “with reasonable means” – instead, IT or statistical experts have to develop tools for these tasks.

The discussion after the presentation showed that it is crucial to influence legislation. Regarding the current draft on the EU General Data Protection Regulation, the most promising way to change something is to lobby national governments by making clear that the main target of the draft are Google and social networks, and not research endeavors. Further, it was noted that it is hardly possible to anonymize biosamples/genomes and that it is instead important to restrict access and to implement elaborate data security concepts.

2.2 BBMRI-ERIC: ELSI Common Services

Jasper Bovenberg presented the perspective of BBMRI-ERIC[2] concerning ELSI issues that should be addressed by common services. The aims of BBMRI-ERIC include:

- Building a dispersed federated infrastructure across Europe
- Enabling sample shipping across borders
- Enabling data transfer across borders
- Addressing harmonization issues (nomenclature, storage, etc.)
- Provisioning of IT, ELSI as well as bio-bank services

As there is no established standard on ELSI for biobanks, the BBMRI ELSI common services will be built by calling the member states (national biobank node representatives as well as national governments) to tackle the following issues together:

[2] http://bbmri-eric.eu/

- Providing concrete advice on ELSI issues
- Providing a help-desk for human sample transfer
- Disseminating results of surveys and studies on ELSI topics
- Developing and Organizing tools/services for ELSI
- Enabling a Common voice of the community
- Providing training/education on ELSI issues
- Checking of ethics compliance of research proposals

The discussion showed that it is not easy to establish such a common service based on a mutual agreement, but BBMRI-ERIC is still in its initial phase and the designation of participating national experts and definition of the services are still in progress. The BBMRI-WIKI[3] could be a platform to distribute useful material to all those involved in establishing the ELSI common service as well as to external users.

2.3 Discussion - What are the main challenges?

The general discussion mainly dealt with the question how much researchers have to know about ethics & legal requirements.

Irene Schluender said that is important not to expect too much legal expertise from the researcher. Tools and services should make researchers aware of the precise ethical & legal requirements and guide them through complex ELSI issues. The most difficult task is to find an entry point for the researcher. Experts such as lawyers should present conclusions to researchers instead of debates. However, the responsibility for making sure that only data necessary for research is used always remains with the researcher.

Jasper Bovenberg disagreed and stated that everyone should know the law. In case of uncertainty, it is up to the researcher to find out how to proceed. The researcher has to be conscious of both, the risk and harms of certain data handling activities. Responsibility for ELSI issues cannot be outsourced; the researcher has a legal and moral obligation to take full responsibility.

[3] http://www.bbmri-wp4.eu/wiki/index.php/Main_Page

Anne-Marie Tassé stated that tools and services should provide a resource to guide researchers concerning how to solve legal and ethical problems. Not everyone can have in-depth knowledge of these issues and a researcher has to rely on the opinion of experts. International/global research and big data are important coming challenges.

Gauthier Chassang emphasized the importance of having tools and services that aid in the self-assessment of the researcher's activities. He proposed a two-part recursive process: scientists should assess the scientific implications of legal requirements presented/digested by the legal experts, and this should guide the legal experts to assess the legal implications of scientific data processing.

Murat Sariyar stated that there is a need for high-level guidance for researchers to enable informed decisions. The researcher is not an expert and cannot anticipate questions that may need to be asked or examined, and therefore needs pre-structured information. General information should be provided before going into details. In complex situations, the researcher should be pointed to experts who can help understand the issues. In conclusion, guidance and raising awareness should be the main goals of tools.

Annemarie Verburg pointed out that legal and ethical issues are tackled in a more or less structured way in clinical studies and that clinical scientists are frequently educated in legal and ethical issues before starting clinical studies. There exists, for example, the 'Good Clinical Practice' and the 'Clinical Trials directive'.

Steffi Suhr responded that basic research might be different, with researchers using sensitive data only in certain research contexts, i.e. to address specific questions that may only come up during the pursuit of a certain line of scientific inquiry. In these cases, no systematic, up-front training is available.

Irene Schluender pointed out that the national requirements are much better known than requirements that relate to international projects due to the differences between countries.

Jaakko Leinonen agreed and pointed out that, for example, data protection requirements vary even between Scandinavian countries.

Murat Sariyar emphasized that it is important to distinguish different use cases: the level of knowledge to be expected from the researcher depends on the type of research, e.g. clinical/medical vs. bioinformatics/basic research.

Anne-Marie Tassé proposed taking a step back to see what a framework that facilitates research should look like in general and what we should work towards building.

The discussion showed that most of the participants do not expect in-depth ELSI knowledge from the researcher, especially when they are involved in basic research. Main challenges that were discussed included:

- ELSI interoperability between nations: each country has different legal and ethical contexts. Therefore, solutions from one country cannot be mapped 1:1 to another country
- Lack of useful solutions: it seems important to gather and prepare existing solutions, and to develop context-dependent use cases
- Dissemination of relevant information: it is not enough to have information and resources available, one should also investigate how the information can be disseminated to the researcher properly

3 Existing resources

3.1 Resource presentations

International Policy interoperability and data Access Clearinghouse (IPAC)

Anne-Marie Tassé presented the 'International Policy interoperability and data Access Clearinghouse' (IPAC)[4] provided by the Public Population Project in Genomics and Society (P3G). P3G is a not-for-profit consortium that provides the international research community with access to expertise, resources and

[4] http://p3g.org/ipac

innovative tools for health and social sciences research. It has over 460 members (institutions/individuals) and seeks interoperability between them by developing networks and support tools. Besides some research programs, including ELSI 2.0, P3G offers the following services under the IPAC umbrella:

- The ELSI Clauses Database[5] is an open access resource allowing users to search and select models of clauses that best suit their needs. The clauses are developed for specific use cases, e.g. rare diseases, and there are no country specific clauses yet. The database, however, does not aim to provide ready-made templates.
- Policy interoperability screening: it is validated whether studies can work together (e.g. consent; confidentiality; etc.). In case of a negative response, development of tools is triggered.
- Data access clearing house: it allows access to controlled databases (e.g. ethics approval/waiver; institutional sign off; etc.) and is centralized as a meta-access committee.

Research groups are approaching P3G to have specific problems and use cases included into the IPAC service portfolio, and one of the most used service is support for REB (Research Ethics Board) approval. There are also many requests for developing suitable consent forms.

Anne-Marie Tassé emphasized that science is global in nature, whereas law and ethics are still operating in silos (differing from country to country).

BBMRI Legal WIKI

Jasper Bovenberg presented the BBMRI Legal WIKI[6], a knowledge platform intended to enhance the embedding of the pan-European Biobanking into the European legal framework. The WIKI provides knowledge and documents/templates, and allows grassroots contributions. It is currently kept up to date under the BBMRI-LPC (Biobanking and Biomolecular Resources

[5] http://www.p3g.org/resources/ipac
[6] http://www.bbmri-wp4.eu/wiki/index.php/Main_Page

Research Infrastructure – Large Prospective Cohorts) project[7] umbrella. A login is required for the user to be able to browse beyond the main pages.

The WIKI is meant to include all information about home state compliance with EU regulations. The information is grouped by nation, and then by topic. National biobank information is to some extent provided in different languages (e.g. Netherlands). In general, the EU standards form a minimum in a country that can be built upon.

It was asked whether it would feasible to make the information machine-readable, i.e. to extract metadata and to annotate the information pieces with controlled vocabularies/ontologies (from the legal field), etc.

Human Sample Exchange Regulation Navigator (hSERN)

Gauthier Chassang presented the Human Sample Exchange Regulation Navigator (hSERN)[8], which provides users with structured information on theoretical and practical legal aspects for exchanging biological samples across borders. The information is presented in the form:

- Overview: provides a general comment related to the selected countries on this topic
- Theory: provides an easy access to different legal notions and to the implemented legal texts, i.e. theoretical information on rules
- Practice: provides an access to the legal or administrative forms and to the actions to undertake, i.e. practical info on what needs to be adhered to in the country of delivery
- Issues: present related questions and relevant documentation

Besides regulations of each country and the corresponding update status on the hSERN website, international standards are given as well. The documents are available as .pdf downloads and as English translation. The user of hSERN can

[7] http://www.bbmri-lpc.org/
[8] http://www.hsern.eu/

contribute to the resource[9] as well as subscribe to updates of the website and its content.

It was emphasized that it is necessary to make the user aware of tools and services such as hSERN. Frequently, they are not found by performing a simple Google search; e.g. googling “biobank ethics requirements” does not bring up any of the tools discussed in the workshop among the top hits.

BioMedBridges Legal & Ethical Assessment Tool

Murat Sariyar presented the BioMedBridges online tool for the assessment of legal and ethical requirements[10]. The tool was developed in order to aid biomedical researchers and raise their awareness with respect to ELSI by providing legal and ethical requirements as well as ways to meet them when making data from different data sources available or using them in new contexts. The requirements relate to the EU-level and fall into four clusters, which were refined and then formed the basis for the tool: data protection, data security, intellectual property and biosample security.

To enable a structured query based on all requirements, they were described using certain attributes. On that basis a workflow was defined and implemented: users are guided by a structured query process during which they are asked about several data characteristics (i.e. the attributes describing the requirements). For example, users are asked about the data category (metadata, text data, images, genetic data, biosamples or biosample associated data), the extent of disclosure (pseudonymous versus anonymous), or the level of use restrictions (e.g. intellectual property requirements). At the end, the tool filters and displays rules, regulations and corresponding solutions that are important for the scenario specified. All background, templates etc. provided by the tool are also available separately online as part of the corresponding BioMedBridges project deliverable[11].

[9] http://www.hsern.eu/index.php/news/show/international-code-of-conduct-for-genomic-and-health-related-data-sharing-public-consultation
[10] http://hhu3.at.xencon.de/web/guest/bmb-assessment-tool
[11] http://www.biomedbridges.eu/deliverables/52-0

Even though the first instance of the tool includes mainly EU-related requirements, the requirement matrix of the tool can be extended to cover specific national requirements (although this may require adjustments of the workflow, especially when new attributes are used to describe the requirements).

It is not easy to ensure that the output is useful without having experts that test the tool. It is therefore important that feedback from such experts as well as from users is enhanced by providing meta-data information and conducting surveys about the tool and its usefulness for researchers.

Resource Entitlement Management System (REMS)

Jaakko Leinonen presented the Resource Entitlement Management System (REMS)[12] developed by the IT center for Science (CSC). REMS is an open source software for managing access rights to research resources, such as research datasets. First, a distinction between different kinds of access to research datasets is made:

0: full public access
1: Researcher has role/group access
2: license terms for use of database
3: Submission for access

REMS aims to assist with access levels 2 and 3 by providing a workflow for enabling electronic submission of data-based authorization policies. Data can be provided together with rules for accessing them. Such rules can be very specific or rather general based on the contexts the data is destined to be used in. Researchers can search the data repository that is governed with REMS via meta-data that is provided together with the datasets. When researchers decide to use a dataset, they apply for access and have to commit to license terms. This application for access is then submitted to an approver, e.g. an access committee, who decides on the basis of certain policies and rules to grant access or not.

[12] https://confluence.csc.fi/display/REMS/Home

3.2 Classification of Resources and Tools

There are many categories for classifying the services presented. One important issue is **maintenance/curation** of the services. For example, with respect to

- IPAC: web resource toolkits and catalogs are updated by experts mainly as P3G projects require verification and approval of further project data accesses or ethics framework
- hSERN: information resources are validated and checked by national lawyers. In view of specific projects/needs, information is updated by experts together with project partners and other contributors. The contributions depend on the goodwill of national experts.
- BBMRI: only validated documents (e.g., approved templates from ethics committees) are uploaded. This means that documents are stemming from approved sources or national editorial sources. However, updates depend on availability of funding. Cross-links between resources are reducing the work for updates.

Other ways to categorize services might include:

- Is the service based on the availability of **legal experts** that can be contacted for consultation? – True for IPAC (data access clearing house), BBMRI-WIKI (part of future combined services) and hSERN
- Is the service an **information resource**? – True for IPAC, BBMRI-WIKI, hSERN and the BMB online tool
- Is the service providing an **automated tool**? – True for IPAC (ELSI Clauses Database), BMB online tool and REMS

In addition to the categorizations above, it is of high importance to have quality metrics that can be used to assess the usefulness of the services. Access statistics are only an indirect quality measure. More than that, one should know how and in which contexts the services are used or whether they facilitate their own improvements. Quality tests with respect to the content depend on legal experts/judges.

4 Mapping tools and resources/requirements

Gauthier Chassang proposed a workflow that would integrate services: the BMB online tool would be used for general information and raising awareness, the BBMRI-WIKI would be used more detailed information as well as templates and finally HSERN as well as REMS would be placed in the phase of concrete exchanges of data or biosamples. IPAC seems to provide a similar integration based on its own services.

In order to map automated tools and information resources the following questions were posed:

Who is the target user group of the service?

- IPAC: researchers doing original collaborations (projects), consortia of researchers
- BBMRI-WIKI: community of biobanks, funders, stakeholders, pharma, patients
- hSERN: researchers (main), lawyer, IT services, ethics committee
- BioMedBridges online tool: researchers

For which purposes could the service be used (e.g. funding, counseling)?

- IPAC: supporting ongoing consultations on legal aspects
- BBMRI: counselling
- hSERN: counselling (mainly informative)
- BioMedBridges online tool: counseling

During which stage of a project is the service intended to be used?

- IPAC: beginning or before start of the project and also ongoing because data access can be at any time following
- BBMRI-WIKI: it varies, can be used at every phase of the whole process
- hSERN: mainly international high level, therefore mainly at the beginning of researches involving human samples and personal data sharing

— BioMedBridges online tool: before data access at project start and during project phase before linking or sharing data.

What level of granularity is provided?

— IPAC: information can be as granular as needed (mainly genomics)
— BBMRI: highly specific (with approved templates), especially fundamental research is covered
— hSERN: covers basic information, but is working to increase granularity - specific samples exchanges (e.g. stem cells, embryos…)
— BioMedBridges online tool: more or less general requirements related to fundamental research

Based on the answers to these questions it was acknowledged that there is indeed a great potential to collaborate as the tools have some overlap while they cannot substitute each other. A possible mapping is from the BioMedBridges online tool ⇔ BBMRI-WIKI: for example, the BioMedBridges online tool could be used to make the information presented in the BBMRI-WIKI more digestible for the user.

A further useful resource that should be considered is the international database on ethical, legal and social issues in human genetics[13].

5 Summary and action items

In summary, the discussions showed (1) that the different approaches assume different user needs and prior knowledge of the subject, affecting how a service is designed and the usefulness of it is assessed, (2) that there is much potential for collaborations, and (3) that enriched annotations of services (e.g. update status, completeness of information, disclaimers etc.) would increase their value and facilitate their quick assessment by (potential) users. Overall, it became obvious that there is still work to do with respect to providing researchers using sensitive data or samples with truly ‘useful’ tools that do not require pre-existing, in-depth

[13] http://www.humgen.org

knowledge of Ethical, Legal and Social Implications (ELSI) or plenty of time to delve into the details.

The following action items were agreed:

- Define a common set of metrics/indicators to assess the usefulness of a service: establish a working group including Anne Marie Tassé, Steffi Suhr, Jasper Bovenberg and an ECRIN representative
- Perform joint surveys to determine and refine user/target groups of the services (it was open whether a working group should be defined)
- Formulate a best practice guideline for the services including concise annotation of the (e.g. update status, completeness of information, disclaimer - what is covered and what is not, where to get additional information/access expertise), provision of information in machine-readable format etc.
- Cross-link services where appropriate, e.g. from the BioMedBridges tool to IPAC and use of resources from successful applications for data access (e.g. good quality templates/legal agreements)
- Set up a mailing list for exchange of information, e.g. service updates, extensions, the status of the BBMRI-ERIC ELSI common service: bms-ethics@elixir-europe.org (initially only to include the participants of the workshop)
- Raise awareness of resources in the research community, e.g. by providing a BioMedBridges training online module that gives an overview of ELSI and provides information on ELSI services as well as what they cover (the annotation of services is an important preliminary work for this task). On that basis, cross-linking between services can be established more easily.